

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2000, 14:54:45 ; Search time 10764.3 Seconds  
(without alignments)  
2150.454 Million cell updates/sec

Title: US-09-358-322-1  
Perfect score: 5300  
Sequence: 1 gatcaataaatagaaccca.....taccatttactactgac 5300

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1033670 seqs, 2183789903 residues

Total number of hits satisfying chosen parameters: 2067340

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba1.\*  
2: gb\_ba2.\*  
3: gb\_om.\*  
4: gb\_ov.\*  
5: gb\_pat.\*  
6: gb\_ph.\*  
7: gb\_pl1.\*  
8: gb\_pl2.\*  
9: gb\_pr1.\*  
10: gb\_pr2.\*  
11: gb\_pr3.\*  
12: gb\_ro.\*  
13: gb\_sy.\*  
14: gb\_un.\*  
15: em\_fun.\*  
16: em\_hum1.\*  
17: em\_hum2.\*  
18: em\_in.\*  
19: em\_om.\*  
20: em\_or.\*  
21: em\_ov.\*  
22: em\_pat.\*  
23: em\_ph.\*  
24: em\_pl.\*  
25: em\_ro.\*  
26: em\_sts.\*  
27: em\_sy.\*  
28: em\_un.\*  
29: em\_v1.\*  
30: gb\_ba3.\*  
31: gb\_in1.\*  
32: gb\_in2.\*  
33: gb\_in3.\*  
34: gb\_pl3.\*  
35: gb\_pr4.\*  
36: em\_ba1.\*  
37: em\_ba2.\*  
38: em\_bt1.\*  
39: em\_bt2.\*  
40: em\_bt3.\*  
41: em\_bt4.\*  
42: em\_bt5.\*  
43: em\_bt6.\*

44: em\_bt7.\*  
45: em\_bt8.\*  
46: em\_bt9.\*  
47: em\_bt10.\*  
48: em\_bt11.\*  
49: em\_bt12.\*  
50: em\_bt13.\*  
51: em\_bt14.\*  
52: em\_bt15.\*  
53: em\_bt16.\*  
54: em\_bt17.\*  
55: em\_bt18.\*  
56: em\_bt19.\*  
57: em\_bt20.\*  
58: em\_bt21.\*  
59: em\_bt22.\*  
60: em\_bt23.\*  
61: em\_hum3.\*  
62: em\_hum4.\*  
63: em\_hum5.\*  
64: em\_hum6.\*  
65: gb\_pr5.\*  
66: gb\_pr6.\*  
67: gb\_pr7.\*  
68: gb\_bt1.\*  
69: gb\_bt2.\*  
70: gb\_bt3.\*  
71: gb\_bt4.\*  
72: gb\_bt5.\*  
73: gb\_bt6.\*  
74: gb\_bt7.\*  
75: gb\_bt8.\*  
76: gb\_bt9.\*  
77: gb\_bt10.\*  
78: gb\_bt11.\*  
79: gb\_bt12.\*  
80: gb\_bt13.\*  
81: gb\_bt14.\*  
82: gb\_bt15.\*  
83: gb\_bt16.\*  
84: gb\_bt17.\*  
85: gb\_bt18.\*  
86: gb\_bt19.\*  
87: gb\_bt20.\*  
88: gb\_bt21.\*  
89: gb\_bt22.\*  
90: gb\_bt23.\*  
91: gb\_sts1.\*  
92: gb\_sts2.\*  
93: gb\_v11.\*  
94: gb\_v12.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	314.6	5.9	1306	2	CRPCS20	X58242 C.ruminanti
2	155.2	2.9	312430	2	RPIX02	AJ235271 Rickettsi
3	149.2	2.8	14001	33	PFCOMPIRB	X95276 P.falciparu
4	144.8	2.7	69034	31	AF007261	AF007261 Reclinomo
5	140.6	2.7	317511	2	CJ11168X3	AL139076 Campyloba
6	138.2	2.6	19517	33	DMU37541	U37541 Drosophila
7	137.6	2.6	4601	33	DMU11584	U11584 Drosophila
8	136.2	2.6	104992	68	AC005504	AC005504 Plasmodiu
9	136.2	2.6	163226	68	AC004157	AC004157 Plasmodiu
10	133.2	2.5	14867	31	AE001398	AE001398 Plasmodiu
11	128.4	2.4	86827	33	PFMAL3P5	AL034556 Plasmodiu
12	124.2	2.3	86827	33	PFMAL3P5	AL034556 Plasmodiu

```
13 123.8 2.3 67970 33 PFMAL1P3 AL031746 Plasmodiu
C 14 123.6 2.3 4601 33 DMU11584 U11584 Drosophila
C 15 123.6 2.3 19517 33 DMU37541 U37541 Drosophila
C 16 120 2.3 161891 68 AC008206 AC008206 Drosophil
C 17 119.4 2.3 104992 68 AC005504 AC005504 Plasmodiu
C 18 119.4 2.3 163226 68 AC004157 AC004157 Plasmodiu
C 19 117.6 2.2 161891 68 AC008206 AC008206 Drosophil
C 20 110.8 2.1 15421 33 PFCOMP1RA X95275 P.falciparu
C 21 110 2.1 75076 9 AC004948 AC004948 Homo sapi
C 22 108 2.0 106650 10 AC007708 AC007708 Homo sapi
C 23 107.8 2.0 67970 33 PFMAL1P3 AL031746 Plasmodiu
C 24 107.6 2.0 1275 66 HSU79270 U79270 Human clone
C 25 107.6 2.0 2676 11 AF044321 AF044321 Homo sapi
C 26 106 2.0 2426 34 SDU49822 U49822 Saccharomyc
C 27 105.8 2.0 234112 33 PFMAL1P2 AL035475 Plasmodiu
C 28 104.4 2.0 128779 65 HS29K1 298745 Human DNA s
C 29 104.4 2.0 145670 10 AC008132 AC008132 Homo sapi
C 30 104.4 2.0 187543 10 AC005678 AC005678 Homo sapi
C 31 100.6 1.9 1496 32 CEY53C12D AL033123 Caenorhab
C 32 100.6 1.9 242893 90 CEY53C12 292859 Caenorhabd
C 33 99.8 1.9 298987 31 AE003846 AE003846 Drosophil
C 34 99.4 1.9 17443 32 CEY52E8A AC024806 Caenorhab
C 35 99.2 1.9 12900 31 AE001429 AE001429 Plasmodiu
C 36 99.2 1.9 122364 10 AC007324 AC007324 Homo sapi
C 37 99.2 1.9 159550 71 AC013360 AC013360 Homo sapi
C 38 99.2 1.9 170102 10 AC008079 AC008079 Homo sapi
C 39 98.8 1.9 321003 90 PFMAL1P3 AL035476 Plasmodiu
C 40 98.4 1.9 8514 34 SPAC1420 AL121732 S.pombe C
C 41 98.4 1.9 94384 11 AC011718 AC011718 Homo sapi
C 42 98.4 1.9 208684 83 AC072048 AC072048 Mus muscu
C 43 96.8 1.8 124760 70 AC011355 AC011355 Homo sapi
C 44 96.8 1.8 173915 35 AP001254 AP001254 Homo sapi
C 45 96.8 1.8 340000 35 AP001678 AP001678 Homo sapi
```

## ALIGNMENTS

```
RESULT 1
CRPCS20 CRPCS20 1306 bp DNA BCT 30-JUN-1993
LOCUS C.ruminantium (clone PCS20) DNA for unknown ORF.
DEFINITION X58242.1 GI:40639
ACCESSION X58242.1 GI:40639
VERSION identified open reading frame.
KEYWORDS heartwater rickettsia.
SOURCE Cowdria ruminantium
ORGANISM Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
REFERENCE 1 (bases 1 to 1306)
AUTHORS McGuire,T.C.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-1991) T.C. McGuire, Washington State University,
Dept. of Vet-Micro/Pathology, Bustad Hall, Pullman WA 99164-7040,
USA
REFERENCE 2 (bases 1 to 1306)
AUTHORS Waghela,S.D., Rurangirwa,F.R., Mahan,S.M., Yunker,C.E.,
Crawford,T.B., Barbet,A.F., Burridge,M.J. and McGuire,T.C.
TITLE A cloned DNA probe identifies Cowdria ruminantium in Amblyomma
variegatum ticks
JOURNAL J. Clin. Microbiol. 29 (11), 2571-2577 (1991)
MEDLINE 92129532
COMMENT See also X58243.
FEATURES Location/Qualifiers
source 1..1306
/organism="Cowdria ruminantium"
/strain="Crystal Springs (zimbabwe)"
/db_xref="taxon:779"
/clone.lib="genomic: E.coli pUC 19"
/clone="PCS20"
717..1178
/note="unknown ORF"
/codon_start=1
CDS.
```

```
/transl_table=11
/protein_id="CAA41197.1"
/db_xref="GI:40640"
/db_xref="SPTREMBL:Q04302"
/translation="MTOKAKNTIYLLILILMLCLVYASVPLYSFCKVTGVGGTVR
TTIVTQSKGKTKIRFNADINKELPKWFYEIPYTTVKPGEQKLIYFRAENLITNEY
VSGMAVYNTPYKGVKFNKACFCFTKOTLSPYOKTIMPV5FFIDQVYIN"
BASE COUNT 533 a 204 c 183 g 386 t
ORIGIN
Query Match 5.9%; Score 314.6; DB 2; Length 1306;
Best Local Similarity 80.5%; Pred. No. 2.4e-28;
Matches 368; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
QY 5 aaataaataaacaagaataaacaactatttgggtaataagtagcaataactatc 64
DB 710 AAATAAATAACACAAAACCAAAATACTATCTATTATTATTAATATCATCTATC 769
QY 65 tatgatatgttttagtgtaacgcttctgtaccacatatatgtatttttggtaaaagtaacag 124
DB 770 TATGTTATGTTTAGTATATGCTTCAGTACCATTATACAGCATATTCGTAAAGTTACAGG 829
QY 125 ttatgagggtacgtaagaacaagtaataatatacaaatcttaaaatagggttaacactattat 184
DB 830 ATACGGAGGAACGTGCAGAACAACTACAGTAACACAATCTAAACTCGGTAAAGACTACCAT 889
QY 185 taaagtcagatttaagtcagatatatacaacaacacgcccaggaattctatccagaagt 244
DB 890 TAAATAAAGATTTAATGACACATAAATAAAGAACCTCCATGGAANAATTTTATCTGAAAT 949
QY 245 atctcatgtatttgaataacccaggagagaacaaaattgattttaccgcgcagaaaaatct 304
DB 950 ACCTTATACAACGTCTCAACACCAGGAGAACAAAATAAATTTTATAGACAGAAAAATTT 1009
QY 305 actgtatgaggacacttcagggaatgctgtatataaagttacacacacataaagtaggaaa 364
DB 1010 AACTAATGAATACGCTGTCAAGGTATGGCTGTATATAATAATGTTACCCCTTACAAAAGTAGTAA 1069
QY 365 atatttaataaggtagctgttttggttccaccaacaacacattacccctcatcaaaa 424
DB 1070 ATATTTCAATAAGTAGCTGCTGCTGCTTTTACCAACAACAGCTTATCTCTTATCAAAA 1129
QY 425 aactataatgccagtagtcatttttttatagatccagcc 461
DB 1130 AACTATTATGCCGTATCATTTTTTTTATAGATCAAGTC 1166
RESULT 2
RPXX02 312430 bp DNA BCT 12-NOV-1998
LOCUS Rickettsia prowazekii strain Madrid E, complete genome; segment
DEFINITION 2/4
ACCESSION AJ235271 AJ235269
VERSION AJ235271.1 GI:3868717
KEYWORDS complete genome.
SOURCE Rickettsia prowazekii.
ORGANISM Rickettsia prowazekii.
Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
Rickettsiaceae; Rickettsia; typhus group.
REFERENCE 1 (bases 1 to 312430)
AUTHORS Andersson,S.G., Zomorodipour,A., Andersson,J.O.,
Sicheritz-Ponten,T., Alsmark,U.C., Podowski,R.M., Naslund,A.K.,
Eriksson,A.S., Winkler,H.H. and Kurland,C.G.
TITLE The genome sequence of Rickettsia prowazekii and the origin of
mitochondria
JOURNAL Nature 396 (6707), 133-140 (1998)
MEDLINE 99039499
REFERENCE 2 (bases 1 to 312430)
AUTHORS Andersson,S.G.E.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1998) S.G.E. Andersson,
Siv.Andersson@molbio.uu.se, Dept. of Molecular Biology, University
```

COMMENT FEATURES source	of Uppsala, Husargatan 3, Uppsala, S-751 24, SWEDEN On Nov 13, 1998 this sequence version replaced gi:3860789.	
	Location/Qualifiers	
	1. .312430	
	/organism="Rickettsia prowazekii" /strain="Madrid F" /db_xref="taxon:782" 76. .1830 /gene="RP226" 76. .1830 /gene="RP226" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="CAAL4699.1" /db_xref="GI:3860789" /translation="MLNKLCDLFLINLLVTVGASPPPLPSPSLPIIAYDVTDKNI STNSISFEKPFQKFSKQKKNISSOHEQOTKAIHOESQIDISRELNEHQSEPEFI DFGSTITLDSVASIDSKAENSTNLAASNTQIQVQKEFPDSEASEPIDIGNTK FTSATNHEMYKEAVSSNDKETNLSTNIITPNVPVISIPTAQDVNYYVPSQSVQIY KPTNLTSIRNPILNHHDDLKNEKNLESTISNTMTPTNNVSPSIQDTIOTTLNIT VPTAEVHSVQTSVTHMSNHOQAQIPITISINPTVETSTVLRTATESVPINNQEIF VSESETKQDWYTPIMPLVVDPNKQSPLALEOKNNNDQIINNQAESHVSSSNV TIQKDNKVNATSESTKEFKVNETQMLFLPDDDIVLGKLTEDATLEQMDHGVKLF QKKEWIANAKRLVESFIKYYDDINKNDIYANLSYSAVDNAFRAVENNNLFELR ALLDYPILQAKNSTGETLLTSSYNGNYILAKFLVIRGITSVLKNDCEKPLDIAL AKGNTNIVCMMLKAKGN"	
gene CDS	2025. .4013 /gene="RP227" 2025. .4013 /gene="RP227" /codon_start=1 /transl_table=11 /product="DNA GYRASE SUBUNIT B (gyrB)" /protein_id="CAAL4690.1" /db_xref="GI:3860790" /translation="MSDLFSLNKEKKKIVITNYSAKDIEVLGDELPVRRKPGMYIG TDSNAMHLVSELDNDAMDEAVGASIIIMIKHODHSITIFDNGRGPIDNHPKFPD KSALEVILTTHSGSKSNVYHTSGGLHGIGISVYNALSKEFKIKVYKQGLKYSQSY SKGAKLTLISAERAKLRGTISNFTDPPEIFSEKLNPKKIVEIARSKAYLYRGVS IEWECEVEVPDIPKALINPGLKDYLSKKISLDNLVPIEIFSGNIESVDDIKLE WAICWQNDTSAPQSCYNTVPTPQGGTHEOGLKSAILRGLKAYSEMIGNKKAANLII EDILETASIVLSIFIVFSPFOGQTKERKLVSGVSKLVENIHKDFHFLSDSKVLATH LLEHVIATAEFRNNKKNRISRSVTKRLPKGLADCTRTSAEGTELFIVEGDSAG GSAQKARNETQAVLPWGLVNLVNASSTLEKIIINQAIDLEIALACGSLKNYKKNEL RYEKIIMTDADVGAHTASLMTFFFLRMPKLYEEHGLAKPLPLTQSNKIYYA CDEEKIKLTVKLSKAKTEVGRFKGLGEMPAQLKETTMMHPEKRSLLKVTLEDVQ NVDKIVDLGSKKEKPEQFIYEQALVKMDQIINKLDI" 4122. .5447 /gene="RP228" 4122. .5447 /gene="RP228" /codon_start=1 /transl_table=11 /product="TAIL-SPECIFIC PROTEASE PRECURSOR (ctp)" /protein_id="CAAL4691.1" /db_xref="GI:3860791" /translation="MLRFLIATFLSINCATGKETENKISQEARVYKQFQDVEERIEK DYVQVPRQKMDAINGMLSLPHSNYYTDELEDIFETTKGEFGGIGVEIWDSC AIKTISSIDDDLPAGKUGGDIYGVNDELSTLGPKNKAKEMRGPPTGTVRLLIIR EEEAQPOLETRIEIKPIKAHLEKNNAIYIRITTFNESTISLKAARVKKLTKESK DNLKGIIILDRNAGGILDQAIYVSDYFIDSGVITVTKGRTTSNSETKANEFSLKAP KMPVLINGNSASASEIVAGALQDHKRAIILGTKSGKGSVQALTOINSRAAVKLTI SKYTPSGRSIQAGEIPDILIEPAKVEYPEVKIDKREFSESLKNLYLKNDAKNKDS NKEKTKNNKQSESELSLYKKIQPARAYDVIYGLIINTNLETQGRAK" 5653. .7146 /gene="RP229" 5653. .7146 /gene="RP229" /codon_start=1 /transl_table=11 /product="HISTIDINE KINASE SENSOR PROTEIN (barA)" /protein_id="CAAL4692.1"	
gene CDS	/db_xref="GI:3860792" /translation="MNNKYYIVRLSFIILFLINIVINMLFYRYFMKEMIKIQAVALHT KIVELYTDNIWNHKNKISYKHKFDYLLKQDQDFINEVKITAQWFTNLNINISYDL KNGKFIITSMMLHSDVNYKDDSLIEIVITKIDIFLKFSSTAAPLAKFGITSYHIL LPKVVIENESDLTIEHASFVTSYIPVIDHNDLPDVAFEINTNITTSOMKNTISLEOK VFIFIIIFIIICIIISNTNAYARQIIIEEQLETRNLKAQIEKVEKTSSTNTKFPANI SHEURPLNALIGFSEILMSRDEKSNYINDHAGKHLHLSINDLUSKASADK LKVDNIDDLNKLSSLLIKPRAYQAEVLEISRPRKHEVVIINADPRKLQVLLNLL SNAYFTNSGGSVTSISLEKDELAKLYIKVITDGTIGIGIEKDIPTKSAFGQIDSELSR KYEGTGLGLPLTKKLVELMNGKFDLQSKINKGTVTITKYDGSIEI" 7374. .8126 /gene="RP230" 7374. .8126 /gene="RP230" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="CAAL4693.1" /db_xref="GI:3860793" /translation="MKLIILLITFLSFLSGKSEMPQGPGLKYAVNNDFENRLDEQE QEIRKLIGKVEILQHKIDLLKQNLNMLNEENIEVLETDDEKFKQDIFDILLEGHND VSKRTFEVNDIAPYKQAYDLAAYDKNKTLEAKDFKFNFIQYPPNLSLSNAYFWY AECFPKQDYGAAINVLKCYKESPKGAKSSDGLLKLALSIGELKKMGKACNILAKLD KEFINTSVSKKMTEDAKIKFGCKINKNNKI" 8123. .8785 /gene="RP231" 8123. .8785 /gene="RP231" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="CAAL4694.1" /db_xref="GI:3860794" /translation="MTDILDEVLDQNEKRLIFFKLVPIIIISIVITIMVIRK NKDNRIKNNQKNGDILVKTIGLDTTKONKALAFNTLENLITSSNTKIKELIALEQVAI RMSKKYLGAKNLLNKIINDNEKYSEISTAYARIAWCCIVLDDQSNIDQKQKLOKYL IFDDEKFPFWATATIIIRALDKHNMKTQAEKNLKLASNNVSDLLQAKQAKALLVS SK" 8812. .10059 /gene="RP232" 8812. .10059 /gene="RP232" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="CAAL4695.1" /db_xref="GI:3860795" /translation="MKKKALLLLPFLISCNGLSGSKSVKNIVDLTPKLVITQNEPIY LDSNTKIYFPNNMLKNKQYSPAKSMIAEPVFIGDMYITLDIRANISAFSEIKNKII WSYNLSKHKDNYIGGGILHNGKLYITYGARLLIVLDAKSGYELIRKELPDIIIRKP IALNDHTILVQTIISNOTIALDSETKTLWDHESIAELISTYSMIPVQHDNVIVTY TGQVLAINKNGEVKNFETNLNDHTAIPNFDTSILCTPVDHSMNLYIATGLKLI KLNLLTGSVLQINADDIOSMLSLIGSLFIINNARQIAALNPETGKVKFVADLNYEK DKRLKSTFLVPFVGVNDNNQORSNLVISVQGLINFNIDNSNGLKMNPHIILKIRNY YGLRSNNTLYFSTDRQVIFGSI" 10279. .10746 /gene="RP233" 10279. .10746 /gene="RP233" /codon_start=1 /transl_table=11 /product="50S RIBOSOMAL PROTEIN L13 (rplM)" /protein_id="CAAL4696.1" /db_xref="GI:3860796" /translation="MKTYSAPSEIEKKWVWDKNNVGLASRAVATMLRGKPKSP TPHLDCCGNIIIIINAEHIKLTGKLNPNADGVYIRHRTGPGGIDKDTTACKLISGYPE RVIKMAVKRMITRNVLGAKQNSNLVYVANCEHPHMAOQPTIYDFASENPKNKK" 10753. .11238 /gene="RP234" 10753. .11238 /gene="RP234" /codon_start=1 /transl_table=11	











```
/note="codon recognized: CCA"
/product="tRNA-Pro"
/anticodon=(pos:3091..3093,aa:Pro)
complement(3054..3127)
/gene="trnP(ugg)"
complement(3140..3211)
/note="trnK(uuu)"
/note="codon recognized: AAA"
/product="tRNA-Lys"
/anticodon=(pos:3178..3180,aa:Lys)
complement(3140..3211)
/gene="trnK(uuu)"
complement(3212..3251)
/note="signal element pal2"
complement(3252..3740)
/gene="nad8"
complement(3252..3740)
/gene="nad8"
/EC_number="1.6.5.3"
/codon_start=1
/product="NADH dehydrogenase, subunit 8"
/protein_id="AAD11860.1"
/db_xref="GI:2258326"
/translacion="MTINKTAQTLFLTELVKGMSLIDYFFRRKKVTLNTPFEKGPUS
PRFGEALRYQTEGECIACKICEAICPAQATIESEPRIDGSRRTYDIDMTC
IYCFCEACPDAIVEGPNFEFATETHEELLYDKEKLQONGRWETEIAANLANEAL
YR"
exon
4098..4135
/gene="trnW(cca)"
/number=1
join(4098..4135,4663..4697)
/gene="trnW(cca)"
/note="codon recognized: UGG"
/product="tRNA-Trp"
/anticodon=(pos:4131..4133,aa:Trp)
4098..4697
/gene="trnW(cca)"
4136..4662
/gene="trnW(cca)"
/note="group 2"
/number=1
4588..4621
/gene="trnW(cca)"
/note="signal element domv"
/note="Region: Signal element domv"
4663..4697
/gene="trnW(cca)"
/number=2
4742..5179
/gene="rp11"
4742..5179
/gene="rp11"
/codon_start=1
/product="ribosomal protein L11"
/protein_id="AAD11861.1"
/db_xref="GI:2258327"
/translacion="MNKIIGFIRIESGRASPPVGPALGRGVNIMOFCKEFNRR
CKOLNKDGVVPVPIITVYDDKTFSEKMKTPSITVYFKHLEISKGASKPKGENFLNI
DIRSLYETAEIKKIDSNDLKSIICKSLIGSVKSGIKTIIVKN"
5207..5881
/gene="rp1"
5207..5881
/gene="rp11"
/codon_start=1
/product="ribosomal protein L1"
/protein_id="AAD11862.1"
/db_xref="GI:2258328"
/translacion="MSIQEFHNVEDAVKLLCYNLNDKNLSDKIIVGFVFNKTSISK
KRVIKQGYTELKGSFSPNPFKGSPEALFTTKLDNFTIDYVGSDDLKIED
NSIKPNLIVLKDVGSLAKYSRVLGPGLMPTPKGTVVENNILLSTLETLLKGSF
RIKYNKYNMQMSVGNLTMTDQILENIKSLLLYLQNLRLPVSRYKTSLEKKIYVTWTHG
PSFIIN"
5963..6544
/gene="rp110"
5963..6544
/gene="rp110"
/codon_start=1
/trans_except=(pos:5963..5965,aa:Met)
/product="ribosomal protein L10"
/protein_id="AAD11863.1"
/db_xref="GI:2258329"
/translacion="MRKHKLIVSEEVVHIFDPFVIVFHYNTVVDSFSLEKKKOE
LOOFIENACIQKNIKOJPFINFKVKNKMKSSLEKSYKNIINLFSPTLLYTTA
NSFICDILIEWKKKNFILLGKYENLVSVADLLEITKISLNKEDEIVKNISFINS
SVVOLTOSLSONQVSIILNLOVLSQKEDSKILL"
6570..11006
/gene="rpoB"
6570..11006
/gene="rpoB"
/codon_start=1
/product="RNA polymerase subunit beta"
/protein_id="AAD11864.1"
/db_xref="GI:2258330"
/translacion="MNEKTPISKKDLRLGRFSSINEIVIPNLLKISKSSYNNFLQ
RFVPSKRLDVGQSLLKLYFGNSFCYSGVSEKNPDECLENGLTYSIPLYVLKVN
MKDNVKTIKYIIGEIPLLTDSGFIINGVENTVINQLYRDSGVYINKSKSPIKTS
LKIRLAGRSIQDLDFYDNFQIKQIDTKLISGNLQNITHQIDKLLDSFKISFRE
IGETFTYSITEYFCYCNMGWIDFNWKEISYRIFKKNKIISTLSLNDLLANDYI
SILTVYNNFISNDKFLNHDFLLDIYKDNKSIFIETENDDDFSYPLELLNKVGSHL
SSFKNSDFLEIENKINYQIFINLKEQPDSTRNSLOVVIIESKKKLLMFFNKSQVH
KDNDFLEKDVLFHNFKLGPIGRLKINLKFGLNEKKELTSIDETRIIRYLIQIDKGT
ENDIYERQSLNRRFNNGIHEHLYSCLSDYLISKEDSKYSKEDYIKNQLFKKINNC
RNAPYFLIRYLKSSCOYTEQANALSYLGHTRKVSIMPGGLKNDLSLEMRDVHIS
TYGICPIETPEKSGSVNNLSITRTNLTFOLETPIYFVKENFVNIIHYLTPOBE
KYKHAHATEVLKDKGSLNTKEYIRYLSLITVDSKNVYIEMSSGOITSASMI
PFLEHNDGNALMGSMOKOAVPLNLEKSLIGTIGETIIGROFFENSKLLKMTN
ENVFGISILKLNKNNIIFSOFVDLFIKNHFLYSTIKHNSNOKTIDIPKYVKNK
INFENNIDNIGYSIKNNELSGNNVLGVGTSNNGHTEFDSIVISERVDNYQSIH
LEVYIMEMEKDSQELIRNPLSFLDVGVIHLGTNVTCTNLDILIKERKQLLSK
NKLIVYDKSIRYKNNKGVVIOIAKFPDSDVLNKHVITTSNLOKSESLLSYIKEE
KIRMLFSLFFIFKNKYKYLGIKKIILDDLNLHNLKNMNSFSFYKTIILYRNN
LINTKLFEDLYKWLICYKYLFSIEIFLAKNKDCLKECIKODNFYKVIDRFKNMYR
EIVKRNKKNDFVLLNKMSSLVFFIDLKKKIPYSMEFLKNFLQEMWILLFLYIK
EYAKQNFVFENDQIFVQKTKSINKKKDCLKNYSLRMVVVVACNKLQVGDKLNG
RYKNGVSRILSLEDMPYLRDGTPLDIVLSPGVSSRMNIGQLLECHLAGVELSK
KLOOIOTIYDIVKSRKYLMLYNNKRELLLSNLSNQICLFLKNLSHGIPSTPIF
NSANNFIDNLRISGLNIOQSVTIIDGITGVLDNKNVTGYVMKMLHLIDHKIHG
RSTGPYNIITQOASKGSRNGORLGEVMEVWALQAYGAAYTLOEMLTFKSDSTSRSS
FLQNFIRGKTSRIDIPRTFKLLRDLRVLGMELNFIVKRSNE"
11093..15247
/gene="rpoC"
11093..15247
/gene="rpoC"
/codon_start=1
/product="RNA polymerase subunit beta"
/protein_id="AAD11865.1"
/db_xref="GI:2258331"
/translacion="MVTKLKNNLSPFKILKHKKKOVKIYINQPIBKUSLDPDEOI
LKSSGEVCTSKTLNKTIEPEPDEGLFCNKFITNSKNNYICNVSRQSPKRCLE
LKDLSKNSIQNLDSMGIHILISPIHPWLLRTKIIHRLAVGNLNIYDYKEIDKILT
DPETIYNYFQMDLLDVFNRLVNLINDFIKKNKNINININDKNIYVAIKKIRI
ITILRLKQRPWEFLKRIPLVSPYRPLAKLDDKLFVADINFLYONIIERNNILE
LLARYSLEKIDRISSLETEIIDEVKKDNPIKGINLOKAVTILMDSFIFGIITPN
KSILIKLSLGGKGRFDILGRINFSGRVTITSGPRHLNOCGLPIELVLELFPFII
LKMRLDNICQTSKLNLTDDYDFKILFNKFTNLNFKLINSIKHPLVLLNRPATLH
RMNIOTFSPLLIYTKAIQLHPLVCSFENADFGDTMAIYIPHTLRISQIESVLMSSST
NIISPTQDEPIISPTODIVFLGYATVMYOGSIGEGRLVGVNEIKNAIINLELDLHS
PIFSKLDKNVLLTSGRIILLYLLEFVNDIFKYVNVKDFDKKEIVSLIKNISLNL
NKTSFILDQIMSGFYAYKAGLSLHKNDINTPNLKFYKTKKKVEDDLKNFSRGI
ISIDERSKRLNENWSKSSILANOILKDVITVPNRKSSIYMLINSARGMSVOMROL
CGRLGMEIKPSGDIILEPIISNFKESGVYDFYNSTHGARKGVITDSLAKTAOSGLYR
KLFPYAKDLFIETEDCGTRKGIIGNLKYQRDGYKIDSFYKNTVGRILVRNKNPYTG
EILVKKGLIVTNLVNLNLIIFGIEFVVVTRSPMTCTRNNGICRKYCGSIGNNNEIVSLE
```

Query Match 2.7%; Score 144.8; DB 31; Length 69034;  
Best Local Similarity 57.7%; Pred. No. 5.2e-09;





```
/note="Cj0670, rpoN, probable RNA polymerase sigma-54 factor, len 416 aa; similar to many e.g. RP54_BACSU RNA polymerase sigma-54 factor (436 aa), fasta scores; opt: 357 z-score: 603.0 E(): 2.7e-26, 28.9% identity in 443 aa overlap, and RP54_ECOLI (477 aa), fasta scores; opt: 365 z-score: 483.2 E(): 1.3e-19, 28.8% identity in 459 aa overlap, 47.4% identity to HP0714. Contains PS00718 Sigma-54 factors family signature 2, Pfam match to entry PR00309 Sigma54_factors, and helix-turn-helix motif at aa 303-324 (score 1227; +3.37 SD)"/codon_start=1/transl_table=11/product="RNA polymerase sigma-54 factor"/protein_id="CAB72947.1"/db_xref="GI:6968132"/translation="MLKQKITQAPKTKISQIIRSLWPLIQANIEDLKENLDRFAEDNPFLNQDSIQTHDKNGYDFYFKHNVSFAVDKSLGKLSKAYVELLEQLPLPFTSKSELAKEITCEDNEEGYFDFHEEFKEYLEIEIRVRARFKFLDGVGKDYKEAFLFALENMELDEIDECFLMIDFENIQNTYKPLEALVLKRFSTPDLFPYFDSRIIVPDIYVYKENGELKVKLINDDYYPEISITQDGLHDFLSHYIKKAKNLVDALMRKATLYKIGLAIYEQYDFMGKEIKPMFTFDFDLALDLERNASTISRAVANKYLSCEGLIPLKDFAFALDEGEISVNGVKEFVANLVKNEIDRNPLSDSKILEIKKEEFKVDIGRITITKYRKHNLNIASTDRKKDYLEG"2656..2679/gene="rpoN"/note="PS00718 Sigma-54 factors family signature 2"2903..4327/gene="dcuB"2903..4327/gene="dcuB"/note="Cj0671, dcuB, probable anaerobic C4-dicarboxylate transporter, len: 474 aa; highly similar to e.g. DCUB_ECOLI anaerobic C4-dicarboxylate transporter (446 aa)/opt: 1280 z-score: 2918.6 E(): 0, 69.1% identity in 466 aa overlap. No Hp ortholog. Also similar to Cj0088, dcuA (37.0% identity in 465 aa overlap)"/codon_start=1/transl_table=11/product="putative anaerobic C4-dicarboxylate transporter"/protein_id="CAB72948.1"/db_xref="GI:6968133"/translation="NDFLTSLEAGQFAIQIIVILICLYFGAKKGIGALGLGGIGILMLVFAPHKPGPAIDVMLTILAVVAGATQASGLDVMQLQIAERILRRNPKEULILAPVFCFTILPGCTGHVVYIMPIIVDIAIKNGIRPERMPVAASISOMGIIASPVS AVSITALLNHNKLACGDGYNLLOITIPSTLFGVLCIGIFSFRGKDLKDRVFOEKLDPKPKYVYGSGLLGVKLKPSKNWAWIFLGAIALVALLGVDFDLRPNWQV VANGIPQVDALGNKMDLSVSVSIOMELLAGSLIIFTKTDARKIGSNEIFKSGMI ALVAVFGISWADMTFVAVHTPMKALGDIVKHPWTVAVMLLISKFVNSQAAISA FVPLALGIGVEPVTVAFAAACYGYYILPTVPSDLATIQFDRSGTHIGKFVNHFSI LPGLIGVITSCIAGYFIAMAAGYL"4469..4657/gene="Cj0672"4469..4657/gene="Cj0672"/note="Cj0672, possible periplasmic protein, len: 62 aa; no Hp match. Contains possible N-terminal signal sequence and second hydrophobic domain"/codon_start=1/transl_table=11/product="putative periplasmic protein"/protein_id="CAB72949.1"/db_xref="GI:6968134"/translation="NRKGYPMDDIIFALLLLLLAFGLSCFFICFCEIQTKDTRNIFLLVL SLILALYFIALLYPERK"4667..6410/gene="kdpA"4667..6410/gene="kdpA"/EC_number="3.6.1.36"/note="Cj0676, kdpA, probable potassium-transporting ATPase A chain pseudogene, len: 1744 bp; highly similar to e.g. ATKA_ECOLI potassium-transporting ATPase A chain (EC 3.6.1.36) (557 aa). No hp match. Note that kdpA (Cj0677) is intact, while kdpC (Cj0678) is a pseudogene"3071..5529/note="G(9)"3071..5529/gene="kdpB"3071..5529/gene="kdpB"/EC_number="3.6.1.36"/note="Cj0677, kdpB, probable potassium-transporting ATPase B chain, len: 681 aa; similar to e.g. ATKA_ECOLI potassium-transporting ATPase B chain (EC 3.6.1.36) (682 aa), fasta scores; opt: 2419 z-score: 2798.9 E(): 0, 56.6% identity in 678 aa overlap. No Hp match. Contains PS00154 E1-E2 ATPases phosphorylation site and Pfam match to entry PF00122 E1-E2_ATPase. Note that kdpA (Cj0676) and kdpC (Cj0678) are pseudogenes"3071..5529/codon_start=1Query Match 2.7%; Score 140.6; DB 2; Length 317511; Best Local Similarity 49.0%; Pred. No. 1e-08; Matches 557; Conservative 0; Mismatches 559; Indels 21; Gaps 6;QY 1093 tattatacaaaagtcggtggaactgatgccagtaggatactctctgattagcacatctttt 1152DB 134900 TTTTATTAAGTGGGTTCAAGAAATGAGATCATGGGAAAAAGTGGCATAGCTCATATGT 134959QY 1153 ttgaacacttaatttagtgaacagaaaaatt-----tctaatctcatcagcacac 1206DB 134960 TGAACATTTTAAATTTTAAAGTACTTAAAAATTTTAAAGCTGGCGAATTTGATGAGATGG 135019QY 1207 ttgataataatagcggaatttcaatgcaagcacatctcaattttgtctatatatactacg 1266DB 135020 TTAAGGTTTGGTGGTGATATGCAAGCACAGGTTTTCGACTATATCTATTATATA 135079QY 1267 aattaataccaaaacaattttatctcttgcaatgattgaatcagacagaagaatgcaga 1326DB 135080 TTAATGTGCGAAGAAAAATTTAGATAAAGCTTTAGAGCTTTTTCGACAACTATATGGCA 135139QY 1327 attttaaggtttaccgacaaagcatataagaagacacaaagtagctcttagaagaagaa 1386DB 135140 ATTAAATTTAAAGATGAAGAATTTACCCCTGAAGAGCAGCTGGTTTGAAGAGCCGTA 135199QY 1387 aaatgag---agttgaagcccaagcaaaaaacatactagaagaagaaatggaaaatgcat 1443DB 135200 GATGGAGAACAGATAATAACCCCTTAGGATATCTTTATTTTAGGCTTTTAAATCATGCAT 135259QY 1444 ttattacaatggatattggcagaccagtagtagtgatgggaacatgaaatagcaactaca 1503DB 135260 TTATGTATCATCTTATCATTTGGACTCCTTATAGGTTTTTTTAAAGGATATAGAAAAATTGGA 135319QY 1504 acaagaagtgtcgaagcctttcataagctacattatagtccttaataatgctatatataa 1563DB 135320 GTATTGAAGATATTAAAGAATTTACAGTATTATTATCAACCTTAAATGCGATTTTAC 135379QY 1564 ttgttaactggagatcagatccacagaagtaatacacacttgcaaaaaacaatactatggga 1623DB 135380 TTGTAAGTGGCGATATAGAAAGTAGGAAGTTTTTGGAGCTTTCAAAAAAACATTTTGAAA 135439QY 1624 aaataccatctaataaagaacccctcaagtcgaagttaggtagaaccacccgcataaaa 1683DB 135440 AGATTTAAAAATACTAAGACTATACCTTAAATTTACACAAAAGAGCCTTAAGCAAGATGGAG 135499QY 1684 caaatatgactttaacattaaagacagttcagtagaataicccagaaactgttttaagt 1743DB 135500 TAAAAAGAAATATATTACATAAAAACTCAGATACACTGAACTATTAGCCCTTGCCT-----T 135553QY 1744 atcaaatcaccaaatggtattaccataaaaaactactacttaacatgatgttagcagaaa 1803
```

Db 135554	ATAAATTCAAATTTCA--AGCACAAAGATATACCTGCTTTAAATGCT-TTATCCGGAGC	135616
QY 1804	tactcggtagtggttaaatcagcctgcttcaaatgatttggtaattacaatccaatag	1863
Db 135611	TTTTAGGAAGTGGGAAAAGTCTTTAATCAGTCGAAATTTTGATTGATAAATTAATTTAA	135670
QY 1864	ttacatcgataaaaaacagattataacttaactgacgacgataattacccttccattg	1923
Db 135671	TCAATGATTATTATGCTTATGTTAAATGATTTGTATAGAT---GAAATTTATTTATTTT	135727
QY 1924	agctatacctaaaaacgggactctcacagaagctgacacagaattcataatgta	1983
Db 135728	TTTGTAAATGTAACCCCAATGTTAAATGCGAGAAAAGTAGAAAAGAGCTTTTAAATAATCA	135787
QY 1984	taaataattatttagaaaaatggaaatttcgacgagaaatttagaaaagtcacaaagataaag	2043
Db 135788	TTGATAAACTTAAATGGGTAAATTTCTCAAAAAGATTTGCAAGAGTGAAAAATAATG	135847
QY 2044	taaaagcacatttaactatgcatttgaacggactcaacttcataatatttttttgca	2103
Db 135848	TCAAAAGTGATTTATTTTCTCACTTAACAATGCTAGTCAGTGGCAAAATATTATTGTT	135907
QY 2104	tgcatctaactacagagaccgctatcagaaaatcagtaattttacgataccatagaca	2163
Db 135908	CATATCTAGCAAGAGAGATATCAATCCCTTTGGTTAAATTTATGAGAAGATATACAAAAT	135967
QY 2164	aagtaagatcccaagattgtaactccgctatggaaaatctttcacaacataaa	2220
Db 135968	TGGAAATTAAGATTTCAITTTCTTGTCGTAATAAATATTATTATCAAGAAAATTCAA	136024
RESULT 6		
DMU37541	19517 bp DNA circular INV 04-APR-2000	
LOCUS	Drosophila melanogaster complete mitochondrial genome.	
DEFINITION	U37541	
ACCESSION	U37541.1	
VERSION	Gr:1166529	
KEYWORDS		
SOURCE	Drosophila melanogaster.	
ORGANISM	Mitochondrion Drosophila melanogaster	
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;	
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 12511 to 12682)	
AUTHORS	Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and Wolstenholme,D.R.	
TITLE	Drosophila mitochondrial DNA: a novel gene order	
JOURNAL	Nucleic Acids Res. 10 (21), 6619-6637 (1982)	
MEDLINE	83090428	
REFERENCE	2 (bases 5269 to 5695)	
AUTHORS	Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.	
TITLE	Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial rRNA genes	
JOURNAL	Nucleic Acids Res. 11 (8), 2411-2425 (1983)	
MEDLINE	8320794	
REFERENCE	3 (bases 404 to 5272)	
AUTHORS	de Bruijn,M.H.	
TITLE	Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code	
JOURNAL	Nature 304 (5923), 234-241 (1983)	
MEDLINE	83245048	
REFERENCE	4 (bases 804 to 1778)	
AUTHORS	Satta,Y., Ishiwa,H. and Chigusa,S.I.	
TITLE	Analysis of nucleotide substitutions of mitochondrial DNAs in Drosophila melanogaster and its sibling species	
JOURNAL	Mol. Biol. Evol. 4 (6), 638-650 (1987)	
MEDLINE	88174373	
REFERENCE	5 (bases 5268 to 13619)	
AUTHORS	Garesse,R.	
TITLE	Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations	
JOURNAL	Genetics 118 (4), 649-663 (1988)	

8212147  
6 (bases 441 to 2967)  
Satta,Y. and Takahata.N.  
Evolution of Drosophila mitochondrial DNA and the history of the  
melanogaster subGroup  
Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)  
91089557  
7 (bases 14215 to 14512)  
Ballard,J.W., Olsen,G.J., Faith,D.P., Odgers,W.A., Rowell,D.M. and  
Atkinson,P.W.  
Evidence from 12S ribosomal RNA sequences that onychophorans are  
modified arthropods  
Science 258 (5086), 1345-1348 (1992)  
93089057  
8 (bases 14917 to 19517)  
Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.  
Sequence, organization, and evolution of the A+T region of  
Drosophila melanogaster mitochondrial DNA  
Mol. Biol. Evol. 11 (3), 523-538 (1994)  
94285822  
9 (bases 1 to 408; 13319 to 19517)  
Lewis,D.L., Farr,C.L. and Kaguni,L.S.  
Drosophila melanogaster mitochondrial DNA: completion of the  
nucleotide sequence and evolutionary comparisons  
Insect Mol. Biol. 4 (4), 263-278 (1995)  
96423163  
10 (bases 1 to 19517)  
Lewis,D.L., Farr,C.L. and Kaguni,L.S.  
Direct Submission  
Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department,  
Michigan State University, East Lansing, MI 48824-1319, USA  
Location/Qualifiers  
1. .19517  
/organism="Drosophila melanogaster"  
/organelle="mitochondrion"  
/db\_xref="taxon:7227"  
/note="derived from new and previously submitted  
sequences; sequence is a composite containing sequences  
obtained from different Drosophila melanogaster strains"  
1. .65  
/product="tRNA-Ile"  
complement(97. .165)  
/product="tRNA-Gln"  
171. .239  
/product="tRNA-Phe"  
240. .1265  
/codon\_start=1  
/transl\_table=5  
/product="NADH dehydrogenase subunit 2"  
/protein\_id="AAC47811.1"  
/db\_xref="GI:1166530"  
/translation="MFNNSKILFIMIGTITVTSNLGAWMGLEINLISFLPDL  
LSDNNLMSTPEASLYFLQVLAFTVLFSILLMLKNNNNINESFTSMITMSALL  
LKSGAPFHEFWFNMEGTMMALMTWQKTAPLWLSYLNKYLILLISVLISL  
GAGGLNQSLRKLMAFSSINHGLWMLSSLMISESWLWILFFPYSELFLVTFMNFIF  
KLPHLNOLFSWKNILKFLTFMNFLSLUGLPPFLFKPLWLVIOQLTLCNQIFMLT  
IMMSTLITFLFYLRICYSAFMNMFENNIMKNNNSINYNMYMTMFTFSIFGLFLIL  
SLPFYMF"  
1264. .1329  
/product="tRNA-Trp"  
complement(1322. .1383)  
/product="tRNA-Cys"  
complement(1403. .1468)  
/product="tRNA-Tyr"  
join(1470. .1472,1474. .3009)  
/codon\_start=1  
/exception="mechanism underlying reading frame shift after  
first codon uncertain"  
/transl\_table=5  
/product="cytochrome c oxidase subunit 1"  
/protein\_id="AAC47812.2"  
/db\_xref="GI:7412849"  
/translation="MSRQWLFSNHHDKIGTLYIFFGAWAGMVGTSLSILIRAEIGHGPH



Db 15543 ATATAATAAATATTCGATTATTAATTAATAATTAACGAATAATAAATAATAATAATAATAA-T 15601  
QY 1330 ttaagggtaccgcaagaagcatttaataagagaaacaaaggtagctcttagaagaagaagaaaa 1389  
Db 15602 TTATTTTAATCACTAAATCTGAATTAATTAATATATATATATATATATATATATATATACA 15661  
QY 1390 tgagagtgtaagccaaagcaaaaacatactacagaagaagaaatggaaaaatgcattttatt 1449  
Db 15662 TATATATATATATATACATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15721  
QY 1450 acaatggatagcgacacagtagtaggagtggaacatgaaatagcaacatacaacaag 1509  
Db 15722 TTATTTATTAATTAATAAATCTTAATAAGTATTTTAAAAAATAAATAAATAAATAAATAA 15781  
QY 1510 aagtgctgaagccttcaaaagcattatagctcctaataatgctatataatgaatgtaa 1569  
Db 15782 ATACTTAATAAATCTATTTTATTAATAAATTAATTTATTAATAAATAAATAAATAAATAA 15841  
QY 1570 ctggagatgcagatccacaagaagtgtaacacacttgcacaaacataactatggaataatc 1529  
Db 15842 AATAAGAAATATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15901  
QY 1630 catctaataagaacaccttcaagctcaagcttagggtagaacaccgcataaaacaata 1689  
Db 15902 ATTTATATATATATATATATA-TAATTTTAATTTCAATTAATAATTAATAAATAAATAA 15959  
QY 1690 tgactttaacatlaaaagacagttcagtagaatacccagaactgttttaagtatcaaa 1749  
Db 15960 TAAAAATATTTTAAATCACAATAATCTGAATAATAATTAATTAATAATATATATATAT 16019  
QY 1750 taccaaatggtataccaataaaacatacatacttaacatgagttagcgaaataactcg 1809  
Db 16020 ATAAAAAATAAGAAATAAATTTATCCCCCTATTCATAAATTTATTTGATATAATTAATACT 16079  
QY 1810 gtagtggtaattcagcctgctttacaatgatttggttaataacaatccaatggtacat 1869  
Db 16080 TAAAAAATATTTTAAAAAATAAATGATTATTAATAATTAATAAATAAATAAATAAATAA 16139  
QY 1870 cgataaaacagattataacttaactgaactgacagcgataatt-----accttccattg 1923  
Db 16140 TTATAATAAATTTTATTAATAAATAAATTTTAAATAAATAAATAAATAAATAAATAAATAA 16199  
QY 1924 aagcttacctcaaaaacgggactctcacagaagctgtagaacagaagaaatccataaatgta 1983  
Db 16200 ATATAAAAAATTTAAAAATGATTTTATAAAAAATTCAAATTCATATATTTATATATATATA 16259  
QY 1984 taataatatttagaaaaatggaaatttcacagcagaataatttagaaagtgcaagtataaag 2043  
Db 16260 CATATAATTTAAATTTCAATTAATAATATATAAGTATAATAAATAAATAAATAAATAA 16319  
QY 2044 taaaagcacattlaacttatgcat-ttagcggaactaacttccatcatcatatttttatggc 2102  
Db 16320 TAAATCTGAATTAATTTGAT 16379  
QY 2103 atgcatctaataactagaggtaccgctatcagaaatcagtaataattttacgataccatagac 2162  
Db 16380 AATTTATCCCCCTAUTCATAAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAA 16439  
QY 2163 aagtagatccaagaatgtaactccgctatggaaaaatactcttcaaaaacaataaaga 2222  
Db 16440 TAAAAAATAAATTTATTTAAATTTATACCTTAATAAATAAATAAATAAATAAATAAATAA 16499  
QY 2223 ttaacgggcatttataccctaagtgagaaatagttatgagaacaatattgtgtacecat 2282  
Db 16500 ATAAATAAATAATTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16554  
QY 2283 taatattgattttcttttccattcaacatcatgcaaat-----gactccaatat 2331  
Db 16555 TAAAAATAATTTTAAAAAATTCAAATTTATTTATTAATAATATATATATATATATATATAT 16614  
QY 2332 taacataaagaagctacaaactaaaaataaaaaatacactatctatagtgtgaacatacaa 2391  
Db 16615 AATTTTCAATTAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16674

QY 2392 cctaccacaatttcccttaaaatttgcattcaagaagaagcaggatcacgcttatgatgcctt 2451  
Db 16675 ATAATTAATTAAT 16734  
QY 2452 tgataagcaaggacttgcatactttcacatacaaaaaattataaacaaggatcaaaaaacaa 2511  
Db 16735 TTCATAAATTTATTTGTATTAATTAATAAATCTTAATAAATAAATAAATAAATAAATAA 16794  
QY 2512 ctatgctccagtttgcacaacaatttagaaggcaaggtat----agacttaaaatttg 2567  
Db 16795 TTATTAATAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16854  
QY 2568 atatagcacctgagacaatttttatatacatcttaaaaccccttcatcagaaaaactttgaagaag 2627  
Db 16855 TTAATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 16914  
QY 2628 cccatagtttactcagtgatgcatattcaacaccgctcacagatcacaagaataatccaata 2687  
Db 16915 ATTCAATTCATATATTTA 16974  
QY 2688 gaataatagcagaacagattgcacatgttaaatcattatattctgcctcctgaatttatag 2747  
Db 16975 GTATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17034  
QY 2748 ctacaacagaaatgaatcacgcttatattcaaaaggcgaccatatattctcaaaaagtttacg 2807  
Db 17035 ATA 17094  
QY 2808 ggacataaatacaacaataataatacaacacaggagaagcgtgc-----attatatata 2860  
Db 17095 TAATTAATAAATCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17154  
QY 2861 aaaaatagtttgcagaaggaacaaatcgttatcagcgagcagagagatgtagatccaaca 2920  
Db 17155 ATAAACTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17214  
QY 2921 cagctacaaatttactagataaataattcttccaaattgcccactctggttaataaaca 2980  
Db 17215 ATTA 17274  
QY 2981 aatacaccacccagatcagcactgtaataagagaagacacatttatattatgtacagagagat 3040  
Db 17275 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17334  
QY 3041 gtaccacaagtgctcaaatgttgcacagacacagtagcaccatcacacagcaagactat 3100  
Db 17335 TAAATAAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17394  
QY 3101 catgcatcaaaacttgcataactatgctaggcggaattgaagctcctcaattccaattattaag 3160  
Db 17395 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17454  
QY 3161 atagaatttaagagaacaggttaggattaaacataccatagtagcaggttcactatccaacatg 3220  
Db 17455 ATAAATTTATATATTTATGAATATTTATATATATATATATATATATATATATATATATAT 17514  
QY 3221 aatca-tagtaatgctctatttggtaacataattcactgataataccacagtagcaaaaatg 3279  
Db 17515 TTTAAATAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17574  
QY 3280 tatatccgtcttcaacagatattatagagcacattaaaaagtagtggtgtgtaggaagcac 3339  
Db 17575 TATTTATATAATAAATCATGTTTTAAAAAATAAACAATAAATAAATAAATAAATAAATAA 17634  
QY 3340 ttttgcattgcacaaatcagtagtatcccaactctttatttttctctctctgttaataacaa 3399  
Db 17635 TTTATATGAATAATATATATTTATTTATTTTATTTTAAAAAATAAATAAATAAATAAATAA 17694  
QY 3400 taatgtagtagagatatgttgaagcttacaaattcacagatcctagatcccgaggttatataa 3459  
Db 17695 TAAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 17754

















Query Match		2.5%;	Score 133.2;	DB 31;	Length 14867;
Best Local Similarity		43.1%;	Pred. No. 1.8e-07;		
Matches 1515;		Conservative	0;	Mismatches 1948;	Indels 53; Gaps 16;
QY	532	aaataaacttcataccggtacattataaactgattaaaaaaataaactatttaattgag	591		
DB	10791	AAGAACCITATTTTATTTTTCGTAATTTTATTTTAAATATAGATGAACGTACAA	10732		
QY	592	caaaataattctattcaacagattctttccaattgagagagattccaaacactaca	651		
DB	10731	TAATTAATTTTTCATGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	10672		
QY	652	actactgctgcaacttctc-----tatacgtgatatacaaaagtgaaataaattaaa	705		
DB	10671	TTGATGTAATTTTAGTTCCTTTTATATATACTATATATATATATATATATATACAT	10612		
QY	706	aactttagttttaagaagaattttataaagttttgaatcaaatatttaattactgata	765		
DB	10611	ATAATGATTCATTTTATTAATAAACAACAAAAAATTCATTAAGAGCTTTTATATAT	10552		
QY	766	taaaataactattaaacattaaacgtcttaattaa-----agtattattatttaact	818		
DB	10551	TATTAATAAATAATCAATTCATTAAGAGTAATTTTAATGAATATAGTAATAATTCITATAT	10492		
QY	819	taatttcataacctttattaaacatttcataataaaaaataacttactctcttttttat	878		
DB	10491	TAATCTTATTTGCTTTTAATTTCAACACGTAAAAAAGAAATAATAATATAGGAATAATG	10432		
QY	879	cacttgattattataaataatcataaaactcccaataaaactatttgcaggttatggt	938		
DB	10431	AAGTTGAATCATAGATTTTCATGAATTGAATTAATAAATAATAGTACTCGTTCGTACCAAA	10372		
QY	939	atgataaatttttactgttttttctatagttttcttcaacaatagccaactcattta	998		
DB	10371	ATTA-AAATAAATAATAGAGTTTATTAATAAATAATGTTAAATTTTGTGAATAAT-ATATA	10314		
QY	999	tccttaacattaaagttcacatgaaattagataatgggaatgggaatgatacacgatt	1058		
DB	10313	GACATCATAAATTTGTATAAATAATATATGATTCATTTATTTATGTTCTCAGATGGTATAGG	10254		
QY	1059	ccaaatcatcgccgaccgagcagtcacatggttattatacaaaagtcggtggaactgat	1118		
DB	10253	AATTTTATTTTATCATATATATTTACATTAAGAATTTTATTAACCTTTTGTAGCGCAA	10194		
QY	1119	gatacagtaggatactcgtgattagcacacttttttgaacacttaagttttagtggaa	1178		
DB	10193	AAA-AAATAGTATAAGCTTAAAAAATAATGTTTCTAAAGGGTTTCCATTTTAAAAATA	10135		
QY	1179	gaaaatttcctcaatctcatcgacacacttagtaatatagcggaatttccaatgcaagc	1238		
DB	10134	GATCATTGATATATTTTCGAAAAAATGTTCTTTTTTTT-GTAGACATTTTTCATACAAATAC	10076		
QY	1239	acatctcaattttgtactatatactacgaatttaatacccaaaacaataatttat-ctctgc	1297		
DB	10075	ACTATTTTTTTTAAAAATGTTTAAATGCTTATGTAAATATATTAATTAATCATTA	10016		
QY	1298	aatggaattgaatcagacagaatcagaatttttaaggttaccgcaagaagcatttaag	1357		
DB	10015	TAATAATAATTAATTCATTTTATCATTTTATGTTCTTTTGTCAATAAAGGGATATATATAT	9956		
QY	1358	agaacaaaggttagcttagaagaagaataatgagagttgaaag-----ccaagcaaa	1411		
DB	9955	ATATATATATATATTTTATAGATATCTTTTAAAGATATATCAACAGTTTCTTTATTTAA	9896		
QY	1412	aaactactagaagaagaattgaaattgcatttttattacattgattgagcagccagt	1471		
DB	9895	AAATGTCCTTTTGCAACTTGAATAATACTAATAATAATGTAATAAATAATAATGATGAT	9836		
QY	1472	agtaggatagggaacatgaaa-----ttagcaactacaacaagaagtgcgtgaagcct	1524		
DB	9835	RACATTTTGAATAATATTAATTCITTTTTTAAATATGTTTCAAAATATTAATTTTACAAGTAG	9776		
QY	1525	ttcataagctacattatagtcctaataatgctatttaattgtaactggagatgcagatc	1584		

DB	9775	TAAATTAATCATGTTACCTTATTTATTCAGAAAAATTTAATTAGAAAAACAATAATAAGTTA	9716		
QY	1585	cacaagaagtaatcacacttcgcaaaataactactatggaataataaccatctaataaaga	1644		
DB	9715	GACATAAAATATTT-----ATGTAAAAAATATTACAGAAATATATATCATTTCTGTAATAATT	9661		
QY	1645	aaccttcaggtcaagtttaggtagaacaccgcgcataaacaataatgactttaaccataa	1704		
DB	9660	AAATTTATTTTATTAATAAATTCATATATCATTTTCAATTCAAAGTTGCTAAATTTTA	9601		
QY	1705	aagcaggttcagtagaatacccgaaactgcttttttaattatcatacacaataatggtatta	1764		
DB	9600	TATATTTTATTAATAAATATATATTTTAAATATAATAATAATAATAATAATAATAATA	9541		
QY	1765	caataaaataactacatacttaacatgctgtagcagaataactcoggtagtggaatacca	1824		
DB	9540	AAAAAATAAATAATATATTTTAAATAATAATACATATATATATTTTAAATTTTAGTA	9481		
QY	1825	gcctgcttacaatgatttgtaattacaatccaatggttacatcgataaaacacagatt	1884		
DB	9480	TTAATTTTATTAATAATATATTTAAAAAAGAAATATATTTTTTTTATAAATAATAATAATA	9421		
QY	1885	ataattacttaactgacgcgcagataattacctttccattggaagctataccttaaaaaacgga	1944		
DB	9420	TAAATATAAATAATAAAGATTAAATAAAAAATAAATTTAAAAATAATATTTTAAATTAATAT	9361		
QY	1945	tctctacgagcgtgtagacaagaatactcaaatgataaataataattatttagaataatg	2004		
DB	9360	ATAATAAATAATTTTATTAATAAATAATTTTAAATAATAATAATAATAATAATAATAATA	9301		
QY	2005	gaatttcagcgaataatttagaaagtcgaagatacaagta--aaagcacatttaactta	2062		
DB	9300	AAATTAATAATTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTAAATAAT	9241		
QY	2063	tgcatctgacgcgaactcttcatactattttttatgcatgcactcaataactactaggagt	2122		
DB	9240	AATATATAAATAATAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	9181		
QY	2123	accgctacagaatacagtaataatttaccgatacacaagaagaagtaagtaaccaagagt	2182		
DB	9180	ACAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	9121		
QY	2183	taactccgctatggaataatctctcaaaacaataataagatttaaccgggcattattacc	2242		
DB	9120	TAAITTAATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	9061		
QY	2243	taatggagaatagttatgagaacaataattgtgtcacacatttaattgattttctttcca	2302		
DB	9060	TATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	9001		
QY	2303	ttcaatacatatgcaaatgactcacaataattacaataaagaagcacaactaaaaataaa	2362		
DB	9000	AATAA-AATAAATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	8942		
QY	2363	atacactatcatattgttgaacatcacactcacacaatttctctaaatttgcatttc	2422		
DB	8941	AAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	8882		
QY	2423	aagaagcaggatgcgcttatgatgccttttataagcgaagcacttgcatcactttcacatca	2482		
DB	8881	AATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	8822		
QY	2483	aaatatataacgaaggtacaaacacactatgctctcagtttgcagcaacaattagaa	2542		
DB	8821	AATATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	8762		
QY	2543	ggcaaggtatagacttaaaatttgatagacacttagacaatttttatcatcataaa	2602		
DB	8761	TAAATAAGTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	8702		
QY	2603	acctta-----tcagaaaaactttgagaagccctagtttttactcagtgattgcatacttcaa	2658		

```

/gene="PFC0575w, MAL3P5.1
/note="PFC0575w (MAL3P5.1), Hypothetical protein, len: 689
aa, possible signal sequence, revised: added new exon 2"

```



```
/codon_start=1
/product="Hypothetical protein, PFC0575w"
/protein_id="CAB38969.2"
/db_xref="GI:7711065"
/db_xref="SPTREMBL:O97258"
/translation="MYLNKVIYIISCFILDFCLSFHLLKMKYKNHNNMKSVTFFL
RSPQITRKFRSRKSNFKKKQKPLFLFNLLKGFSGFGRNAYDOKYIDDIS
NINNTIRIKQVTHKKSKNEFTKENIKQLLHCVFSKIDFKIINNLSYIIKKHQMNSIT
VHSILNOISEKWEKDAENYALHLELLKDNITLFSMMHIMDFEKSQKVICIRD
IKSKKRNKLSIYINLEICTLYETCMCLLIKIKYISHLCLFEFFECFECYCNILR
IYECVGDILRIERINYLCEKKIKFKHMDAIIKMEINMKKDDDLFYNYHDELLRC
FTMKINERNKNNIIRSYNDINNDISIDKQMMNPNIDVINNNISLDEKIEQEPEN
DDENLEKIDYEQFLQNDNTIKYIEEDQPLINDNMSINNNNNITMKKKKIK
DTYNDDYDYKEEDLYQKNIDDIYKNTIGMKSLEEFKNQIEQADIEQFQFLS
NYNLQHGVRKSDNDEKTEHINKNTIKNGYDTLQNMENNFIKKNIDNNISND
ISNNIINIKKLKQSDENLNTSDLIYERLTKVLYIQKIEYLKFKYQVDIINEQ
YPIIRNEKIVDLLNIGIKVMSPDVNSLFEKIKIDSIPNEKORNNQMNKSNKI
IIMFNVM"
misc_feature
604..609
/feature="predicted splice donor sequence for exon 1 of
PFC0575w (revised)"
misc_feature
816..825
/feature="predicted splice acceptor sequence for exon 2
(revised of PFC0575w)"
misc_feature
1024..1029
/feature="predicted splice donor sequence for exon 2 of
PFC0575w (revised)"
gene
complement(3354..6644)
/gene="PFC0580c, MAL3P5.2"
CDS
complement(3354..6644)
/gene="PFC0580c, MAL3P5.2"
/feature="PFC0580c (MAL3P5.2), Hypothetical protein, len:
1097 aa, possible signal sequence, predicted using
hexExon"
/codon_start=1
/product="Hypothetical protein, PFC0580c"
/protein_id="CAB38968.1"
/db_xref="GI:4493932"
/db_xref="SPTREMBL:O97257"
/translation="WGLKRVFVLLYSVPSLFAFLKHKDNENNYETLINNDIE
KIKRIHKNKCSIPLLEFLNYDSIYKNIKLWLYFKFRKKRKEEYIYINVRK
KREAIKYNFDEQNLFNKFIYIYLSYKYLSPHLSLYILKRISEHCVNIP
SLIYKLDNKLNLINERKILYKQINNEHTQOAPTHTHHNNNNKKPLDINHSC
KNTNTSSYTNMKEKINIDYKNIHNFYETKSISYKDNCOHITMLNLLQNTY
DNICRGNTNININFYMIILAYICVKNMEIILLNAYHIEDMKKIKNKNNTS
LFKIYISFFFKKEENHYLDLEDQMHLKKENDKFNYSNENTHNNIKYISDNV
FYDHNSNSCFNKKQKQDDNTKHLIMGKIKPKNKSHKNNKTCNINIER
DOKDILKIFLKGKLDIOILNELYVIMYIRLLFCSLKLISIKKNIHLKKME
FDKNKIYLSADYNNLRNLIRLFRSKNEERENINSFAPFLSKNIYFEDEIG
RSRDNIVNVYDKETNKTNTTNNNNNDNICSNNNDHICSNNDHICSNNDHICSN
NNNICSNNNNICSNNNNICSNNKMLDEFCODNKFNDYTRKKERKRI
YELAKIYNIFDYLGKKEKQEDNTINLYIKKPPWIEFLAKNIIKNDKTSFIEH
NNIYNGDKNNNIIKFKKINLFESSIISFYIKDIYEYNKRLYIYDNLKQFCE
YELKNEHNRKLYKKRAFHYIYNFDOFIINNYHIIHKKNIHIIHLKQCKDE
IDVFKDLYCMNNINNNIYFIYHVDHNECVYRIFKAYKNILLEYNLNEKENIY
YKMKLYLTNNIISNDLYPYSYKNIYQNKYKRNKSFHIFSLKNDIHLLE
LYTORIONCDDIFSIVYKYNPEKNPFLNLYVELHYIVYSEKKKKKFFSISP
MVCVGPKT"
misc_feature
join(7669..7717,800..7829,7912..7940,8064..8106,
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
9369..9505,9613..9838)
/gene="PFC0581w"
CDS
join(7669..7717,800..7829,7912..7940,8064..8106,
8302..8368,8568..8641,8713..8812,8924..9006,9122..9258,
9369..9505,9613..9838)
/gene="PFC0581w"
/feature="PFC0581w, hypothetical protein, len: 324 aa,
revised: new gene prediction, splicing prediction very
tentative"
/codon_start=1
/product="Hypothetical protein, PFC0581w"
/protein_id="CAB90285.1"
/db_xref="GI:7711066"
/translation="MKKKKHYSIKFVNVKKNYKVLCKSDYIRIINYMIGLHIFRL
OKIHIYFANQIASSPFESHKNVYRTKEDIIDGIEKCFWNTIDYLSIESIKODNDFS
NDIKTIVTAMKNDOLLTTSYNNKIDTVNASQWAOQSPVIFLINIKFSHRWSPGA
LKVDEEIVSKNNFSSALSNDNSVTKKIYVDLTLLDNIIESTKTKNFASVGVV
TLKKEKKIWRNLKLSKEKIPNMQVMDMKERHIIITVFINLFFLLSLSHRYHDSV
ONFLKEENKNSDKLODDIDEDEKYFDEILREAKKSEEDYDKDDEEL"
misc_feature
7790..7799
/feature="potential splice acceptor sequence for exon 2 of
PFC0581w"
misc_feature
7830..7835
/feature="potential splice donor sequence for exon 2 of
PFC0581w"
misc_feature
7901..7911
/feature="potential splice acceptor sequence for exon 3 of
PFC0581w"
misc_feature
7941..7947
/feature="potential splice donor sequence for exon 3 of
PFC0581w"
misc_feature
8055..8063
/feature="potential splice acceptor sequence for exon 4 of
PFC0581w"
misc_feature
8107..8112
/feature="potential splice donor sequence for exon 4 of
PFC0581w"
misc_feature
8293..8301
/feature="potential splice acceptor sequence for exon 5 of
PFC0581w"
misc_feature
8369..8374
/feature="potential splice donor sequence for exon 5 of
PFC0581w"
misc_feature
8559..8567
/feature="potential splice acceptor sequence for exon 6 of
PFC0581w"
misc_feature
8642..8647
/feature="potential splice donor sequence for exon 6 of
PFC0581w"
misc_feature
8704..8712
/feature="potential splice acceptor sequence for exon 7 of
PFC0581w"
misc_feature
8813..8818
/feature="potential splice donor sequence for exon 7 of
PFC0581w"
misc_feature
8913..8923
/feature="potential splice acceptor sequence for exon 8 of
PFC0581w"
misc_feature
9007..9012
/feature="potential splice donor sequence for exon 8 of
PFC0581w"
misc_feature
9112..9121
/feature="potential splice acceptor sequence for exon 9 of
PFC0581w"
misc_feature
9259..9264
/feature="potential splice donor sequence for exon 9 of
PFC0581w"
misc_feature
9361..9368
/feature="potential splice acceptor sequence for exon 10 of
PFC0581w"
misc_feature
9506..9511
/feature="potential splice donor sequence for exon 10 of
PFC0581w"
misc_feature
9601..9612
/feature="potential splice acceptor sequence for exon 11 of
PFC0581w"
gene
complement(join(10024..10044,10320..10336,10397..10445,
10613..10674,10773..10798,10929..10999,11157..11237,
11453..11515,11715..11767,11950..12040))
/gene="PFC0582c"
CDS
complement(join(10024..10044,10320..10336,10397..10445,
10613..10674,10773..10798,10929..10999,11157..11237,
11453..11515,11715..11767,11950..12040))
/gene="PFC0582c"
/feature="PFC0582c, hypothetical protein, len: 177 aa,
```







misc_feature	PFC0581w" 7830..7835 /note="potential splice donor sequence for exon 2 of PFC0581w"
misc_feature	7901..7911 /note="potential splice acceptor sequence for exon 3 of PFC0581w"
misc_feature	7941..7947 /note="potential splice donor sequence for exon 3 of PFC0581w"
misc_feature	8055..8063 /note="potential splice acceptor sequence for exon 4 of PFC0581w"
misc_feature	8107..8112 /note="potential splice donor sequence for exon 4 of PFC0581w"
misc_feature	8293..8301 /note="potential splice acceptor sequence for exon 5 of PFC0581w"
misc_feature	8369..8374 /note="potential splice donor sequence for exon 5 of PFC0581w"
misc_feature	8559..8567 /note="potential splice acceptor sequence for exon 6 of PFC0581w"
misc_feature	8642..8647 /note="potential splice donor sequence for exon 6 of PFC0581w"
misc_feature	8704..8712 /note="potential splice acceptor sequence for exon 7 of PFC0581w"
misc_feature	8813..8818 /note="potential splice donor sequence for exon 7 of PFC0581w"
misc_feature	8913..8923 /note="potential splice acceptor sequence for exon 8 of PFC0581w"
misc_feature	9007..9012 /note="potential splice donor sequence for exon 8 of PFC0581w"
misc_feature	9112..9121 /note="potential splice acceptor sequence for exon 9 of PFC0581w"
misc_feature	9259..9264 /note="potential splice donor sequence for exon 9 of PFC0581w"
misc_feature	9361..9368 /note="potential splice acceptor sequence for exon 10 of PFC0581w"
misc_feature	9506..9511 /note="potential splice donor sequence for exon 10 of PFC0581w"
misc_feature	9601..9612 /note="potential splice acceptor sequence for exon 11 of PFC0581w"
gene	complement(join(10024..10044,10320..10336,10397..10445,10613..10674,10773..10798,10929..10999,11157..11237,11453..11515,11715..11767,11950..12040)) /gene="PFC0582c"
CDS	complement(join(10024..10044,10320..10336,10397..10445,10613..10674,10773..10798,10929..10999,11157..11237,11453..11515,11715..11767,11950..12040)) /gene="PFC0582c" /note="PFC0582c, hypothetical protein, len: 177 aa, revised: new gene prediction" /codon_start=1 /product="hypothetical protein, PFC0582c" /protein_id="CAB90286.1" /db_xref="GI:7711067"

Query Match 2.3%; Score 124.2; DB 33; Length 86827;  
Best Local Similarity 42.6%; Pred. NO. 1.2e-06;  
Matches 1718; Conservative 0; Mismatches 2358; Indels 59; Gaps 18;

Qy 1074 ccagcagtcgacatggtattatacaaaagtcggtggaaactgatgcacagtaggatac 1133  
Db 38158 ACGATTTAAAGTTTAAAGAAATTTACTATATTTTAAATATAATTTTAAATTAAC 38217  
Qy 1134 tctggattagcacatttttttgaacacttaagtgttagtggaacagaaaaatttccttaat 1193  
Db 38218 TTTTTCATAAAAGAATATATAGTGTGAATATGCAAAAAAACAACAAAATGSAATTT 38277  
Qy 1194 ctcacagcacacttagtaatataggcggaatttcaatgcaagcacactctcaatttgt 1253  
Db 38278 CATATATTATAACATAATGATAAGTTTAAATATAGGAACCCGATTTATATAAAGATTT 38337  
Qy 1254 actatactacgaattaataaccaaaacaattattctctcttgcaatggatattga--at 1311  
Db 38338 CATGTAATAATCTCCTAATAATGAACAATGCTCTTGTCTTTTGTGACATGTCATCAA 38397  
Qy 1312 cagacegaatcagaatttttaaggttaccgcgaagcatttaataagggcaacaaaggtag 1371  
Db 38398 ATAAATAAAGAATATTTTATAATATAAAAAAATGATATACACCATCATGATAAANA 38457  
Qy 1372 tcttagaagaagaataatgagattgaaagcccaagcaaaaaacatactagaagaagaa 1431  
Db 38458 ATATAACTTAATAATAATATTTTATACAAAGACAGCATATATAAAGGGAAATGA 38517  
Qy 1432 tggaaaatgcattttattacaaatggatgagcagaccagtagtaggagtggaacatgaaa 1491  
Db 38518 ATATAATGAACCTTTCATTAATCTTTAAGTATTTATATAATTTATGTTTTACCAACAGC 38577  
Qy 1492 ttagaacactaca-acaaagaagttgctgaagcccttcaatgaagctacattatagtcctaat 1550  
Db 38578 ACAAAAATCATACACATCTTTATGTATTATGTAATGAATAAGAAAATCAAAAAGTACTTTTACA 38637  
Qy 1551 aatgctataattaatgtaactgagagtcagatccacaagaagtaatacacacttgcaaaa 1610  
Db 38638 AAGATAAAGGAAAAAATAAAAAATTTATTTATTCATTAATAGTTTTTCACATTCAAAAGA 38697  
Qy 1611 caatcactg9gaaataccatctaataataagaacacttcaagctgaagttgaggtaga 1670  
Db 38698 AAAATTACTTAAATTTATGTTACCCCTTAATGTTAAATTCATTAAGTTTACAAAATAAATA 38757  
Qy 1671 ccacgcataaacaataatgacttttaacataaaagaacagttcagtagagaatcccagaa 1730  
Db 38758 TATATAGTATAAAGAAATATTAAGAACATATATTTATATATACATTTTATTTTAAAT 38817  
Qy 1731 ctgttttaagtatcacaataccaaatggattaccacaataaaaaactacatacttaaacatg 1790  
Db 38818 TATTAAATATAAATAATTTTAAATATACAAATATAAATCAATGCTTTTAAATAAATA 38877  
Qy 1791 atgttagcagaaataactcggtagtggtgtaattccagccctgctttcaaatgattggtaatt 1850  
Db 38878 TAAATTTTATTAATATAAATAATTAAGTATTTATATGTTATATAATACATTTTATTTTA 38937  
Qy 1851 aacaatccaatagttacatcgataaaaaacagattataattacttaactgacagcgagataat 1910  
Db 38938 ATTCCTTTCTAATTTAATATGATAACATATTTTATAATTAATTTTACACATTTATTAATA 38997  
Qy 1911 tacctttccattgagctcat--accataaaacgggatctctacagaagctgtagaacag 1968  
Db 38998 AACAATATATAAATAATTAATTTTATTTTAAATATAAATAAATAAATAATTAATAA 39057  
Qy 1969 aaattcataaattgtataaataatttagaanaattggaatttcagcagaataatttagaaa 2028  
Db 39058 TAAATAAAAAAATAATATGTAATAGTTATATATAAATAATTAATTAATAAATAAATA 39117  
Qy 2029 gtgcaagataaagaataaagcacatttaactatgctatttgacggagacttaacttcatat 2088  
Db 39118 TAAATAATAAATAAATAATATAATATAATTAATAAATAAATAAATAAATAAATA 39177  
Qy 2089 catattttatggcgtcatctactacagtaggagtagccgctacaggaatcagtaatttt 2148  
Db 39178 AATTATTTTAAATTAATAAATAAATAAATAATATTTATTTAATAAATAAATAAATA 39237  
Qy 2149 acgataccatagacaaagtaagtatcccaagatgtttaactcgctatggaaaatactttc 2208

Db 39238 ATATATTATTATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39297  
Qy 2209 aaacaataataagataaaccggcatttattacctaataatgagagatagttatgagaacat 2268  
Db 39298 TTATATATATATTTATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39357  
Qy 2269 attggttacacatttaattgattttcttttccatcaatacatatgcaaatgatctcaa 2328  
Db 39358 AATAAATTAATATATATTTATTTAAATAAATAAATAAATAAATAAATAAATAAATA 39417  
Qy 2329 tattaacataaagaagctcacactaaataaaaaatacaactatctatattgttgaacatca 2388  
Db 39418 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39476  
Qy 2389 taacctaccacaacttcttcaaaattgcatccaagaagcaggtacacgtctcatgctgc 2448  
Db 39477 TTATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39536  
Qy 2449 ctttgataaagaaggacttgcatactttacatcaaaaatttaaaaggaaggtacaaa--a 2506  
Db 39537 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39596  
Qy 2507 acaactatgctctcagcttttgcaacaacttagaagggcaaggtatagacttaaaattt 2566  
Db 39597 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39656  
Qy 2567 gatagacactagacaatttttatatatcatcttaaaaccccttatcagaaaactttgagaa 2626  
Db 39657 ATAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39716  
Qy 2627 gccctagttttactcagtgattgcataattcaacacgcgcagatcatcaagaataattcaat 2686  
Db 39717 TCAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39776  
Qy 2687 agaataatagcagaacagattgacatgttaaatcattatattctgctcctggaatttata 2746  
Db 39777 ATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39836  
Qy 2747 gctacaacagaatgaatcagctcatatttcaaggggcaccocatatttcaacaagaatttac 2806  
Db 39837 ATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 39895  
Qy 2807 gggacatttaatacaatacaataatcaacacaggaagacggttcgcttatataataaaaaat 2866  
Db 39896 AATAATTAATAAATTTTAAATTCATATATAAATAAATAAATAAATAAATAAATAAATA 39955  
Qy 2867 agtttgacaaggaacaaatcggttatcagcgacgagagatgtagatccaacacagcta 2926  
Db 39956 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 40012  
Qy 2927 tcaaatctactagataaataatattcttccaaattgcatctggtgaatacaacaataacc 2986  
Db 40013 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 40072  
Qy 2987 ataccagatacagctgttaataatagagaagacacattatattatgacagagagatgaccca 3046  
Db 40073 ATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 40132  
Qy 3047 caaagtgtcataatgttgcacagacacagctaccatatacacagcaaga----- 3096  
Db 40133 TGAACATATTTATTTTAAATTTGAATATGAATAAATAAATAAATAAATAAATAAATA 40192  
Qy 3097 --ctatcatgcatacaactgttccaactactatgctagcggtattgaagtcctcaattcaata 3154  
Db 40193 ATATATTACACACAGTATTATTAAATACTGTTTTAAATTTAAATTTAAATTTAAATTA 40252  
Qy 3155 ttaatgataagaattaa---gagacaagtaggattaaacataccactagtagcagttcacta 3211  
Db 40253 ATTTATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 40312  
Qy 3212 tctaacatgaatcatagtaattgctctattgggtacaattattcactgataataaccacagta 3271



```

DYIINFIKSTKKMEKDLNERSLPNNIYNIMSDVPSVTFFVTCINLNFVFKIF
MSFYVPHIIGNSVGIAIWLSTALYSAMILFEFLPSLFSKSVLIYRDKRIDNMHVL
KEFKLKMNFNSFAKYNINPMKEMKCYCKIPLSLNSIGVFISSISDPIVEVWIFI
YLDKRLNKEEKIETSIIMPLYIKLISNVANFPLNVANMEGIVNIRLNNYINDH
LKYNDKINRYNRYNIDYNTIFLONENITSHDDGTSHNLKHLKNYIKNLTN
MFXYFFYHKMYHKNYINQILSGLLKNDONTNKKICFOEHSNSTYNSSHIHE
KREYENIANSNSTMSNEFEKKKNVEYIKLENGSGLSDYNDKDNHDHILKNFN
LKNSAIILGNVSGCSAFFHSILGDFNHTGNLYIENFEKKMPLYVPQNSWLFMG
NIRKSMILFNGEYNPLYIKYTILOSLLNLSLSTIEHGMKYINDHNLKSKQVRICLA
RALYEHIMHKLCUETKELIOPNEILDRLNNKINSYNNKSKLNNYINPNNEN
YLOKMDNDNFYLLDDIFTSLDPSISKKIFSNLFCINDKDFONCSYFISMMKS
TLDNFIEDLNDVQEVNAIFEQDRLKYRGINSIEMKNNLITKESHWGYSNLT
IDYTRKLFDEVELNVKSNMYKEAVFVKNTESVFEIDSINKYIKMKKKNY
KEHKMKNNKNNNNNSKODHINIMNDNHRNYNDNLGPNSTDSDPTVEDSSS
YLDYTTNSNDEEIVPKYDTHFEFNKSSMPVVKSSNNINPNPSFKYEDNNS
FSGISLEYLTWFOQGVFLVTSVFMILSITDEIRFVFLTMSIISKNKSHSD
TILQKVRLEYEVLPIVLSLTSVSGICFMSIYGNITSALKVHNNTLYSILNAPLYF
YNNLGNIRYELVIDISADYGFRLKYKAFIFRCLISLLIYIMRDCIFIPFY
IILYFVFKRFRSGCKEQRLYLSCHTPLCNIYSNALGKNIIYIKLVKKNYVRI
HYNNRISVFFKWLINWASLYKIFILLTTYIIMHPLYASGIIKLVKKNYVRI
LSTLGCISFSARGVLILFLLDYTHIEKEMCCVORLEFAKISKENASMKENEL
NVITQYKEKNEISDKISAIVEYNLSSUSINSQODESKKKYGIKFNYYVSYK
KITVNGRYKYIDEPSLKNIMYALKNOKIGVKGAGKSTILLISILGLNLSQG
GINLGNLDYRYNRKGEDSIIIGLAOSSFYVYNWNTITFDYNNNTDDEIVHALYN
RHKYKILIDEIPIENLNSVHDELNSFLGKAKSFNYIIRNHPNNTVLIISHANT
LSCDDYIYLRKEIYRCSEYDVKTQSELSHLEMDD"
23896..31533
/gene="rRNA"
/region="region containing small subunit, 5.8S and large
subunit rRNA genes and spacer regions"
23896..31533
/gene="rRNA"
complement(join(31966..32476,32675..32775))
/gene="MAL1P3.04"
complement(join(31966..32476,32675..32775))
/gene="MAL1P3.04"
/region="MAL1P3.04, conserved hypothetical membrane protein,
len: 203 aa, similarity: P. falciparum chromosome 2,
PFB0110W, O96126 predicted integral membrane protein (255
aa), fasta scores: Opt: 335, E(): 4.9e-15, (36.1% identity
in 191 aa overlap)"
/codon_start=1
/product="conserved hypothetical membrane protein,
MAL1P3.04"
/protein_id="CAB63559.1"
/db_xref="GI:6594247"
/translation="MKSYTFINVTILLFTLLFLYYNYDTFSKTFNNNIKIDIN
RFKRIIAESEQKYPWEEDFCILNEEELIRPEHNDSPYLPEHYENIDKINELSINS
TKIKWETIKKRONYKERTDMNHRDPMWHYKWNILYKVKHKNITLKDITNPDI
HDKRETTITWKIKWQIEDIEVLENLQVLEWLRILTLELFYKKE"
complement(32477..32486)
/region="potential splice acceptor sequence"
/region="potential splice acceptor sequence"
complement(32669..32674)
/region="potential splice donor sequence, aaa/gtatat"
join(36657..36743,36864..37343)
/gene="MAL1P3.05"
/region="MAL1P3.05"
join(36657..36743,36864..37343)
/gene="MAL1P3.05"
/region="MAL1P3.05, hypothetical protein, len: 188 aa"
/codon_start=1
/product="hypothetical protein, MAL1P3.05"
/protein_id="CAB63560.1"
/db_xref="GI:6594248"
/translation="MRKMSGIFFKILLICISFICVFECFNKCMISYRKDLLWYSFN
CFNYSIDRSLAEGSESKETKVDIPNIELLSKINYEYKRMKEIVGFMNDNNLN
IANVLIKNHSFTNIENIFSLINDSSKSPVLTKLFKEGFSIFPHMLNNVKKLFLDLCQ
RNPLHILGLVILAAIYVFENKFNFC"
36744..36749
/region="potential splice donor sequence, aag/gtatga"
/region="potential splice acceptor sequence"
36854..36863
/region="potential splice acceptor sequence"
complement(join(38049..39995,40210..40284))

```

```

/gene="garp"
complement(join(38049..39995,40210..40284))
/region="garp"
/region="MAL1P3.06, gap, len: 673 aa, similarity: almost
identical to GARP_PLAFF (678 aa), fasta scores: 97.6%
identity in 678 aa overlap"
/codon_start=1
/product="hypothetical gap protein"
/protein_id="CAB63561.1"
/db_xref="GI:6594249"
/translation="MNVLFSLNYICILFEVVCTLNFSTKCFNSGLNKNQNLNKSFDSP
TNRLLNETEKNKDNDSKSTLLKDEKEDDDVPTTNDNKNVNNNETSSDTPD
TNIIIVNDKENSVDKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
DKKENSEVMSLYTGQHPKNATEHGEENLYEEMVINNNAGGOLLSPYOIRE
QGGCGIISVTHSNDTKDNENISEDKEDHQOEMLKTLDKKERQKEMKEQRE
KTEKKKQOEKKEKQKKEKQKKEKQKKEKQKKEKQKKEKQKKEKQKKEKQK
HKNEEEMQOQDOTSEETNNIEMVPLSPLDVTTPEEHKEGHEHKEGHEG
EHKEEHEHKEHKEHKEHKEHKEHKEHKEHKEHKEHKEHKEHKEHKEHKEH
DKDGEIINLEKDEACEOHITVSRPSQPOCKLIDPEQLTLMOKSVEKNLSIQ
EQLIGTIGRVNVPDRDNHKKMAIEAELQKQHVDEEDKKEEKEVEEESKEVQ
EDEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEEVEE
DDADEDDDDADEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
DDADEDDDDADEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
complement(39996..40005)
/region="potential splice acceptor sequence"
/region="potential splice donor sequence, aag/gtaaca"
join(45401..46396,46562..50233)
/gene="MAL1P3.07"
/region="MAL1P3.07"
join(45401..46396,46562..50233)
/gene="MAL1P3.07"
/region="MAL1P3.07, hypothetical protein, len: 1555 aa"
/codon_start=1
/product="hypothetical protein, MAL1P3.07"
/protein_id="CAB63562.1"
/db_xref="GI:6594250"
/translation="MEQIGIANNIETIGNIVKRSPPFIWRKKFIEVSKKINSLS
NENIECKNCKRKINWDDDEEVEFSKFFVEYISNVKMEFEKAIINTLNKKT
KTCLRSKKYPKEIRNIMGEIDYIDAINDKQFANLYCWSERIDYKWLNEKMEY
Query Match 2.3%; Score 123.8; DB 33; Length 67970;
Best Local Similarity 43.5%; Pred. No. 1.5e-06;
Matches 935; Conservative 0; Mismatches 1187; Indels 29; Gaps 7;
Qy 490 taaactcatcactcttcctatgtatctttaaagtaacaaagaataaactcattaccg 549
Db 7823 TATAAATAATTCCTTAATATATATTAACATTTATAAATAATAATTAAGGAATAAT 7882
Qy 550 tacattataactgattaaaaataactatttaattgagcaaaataatttatctatt 609
Db 7883 TTTCTTAAAAACGATGATATAAATAATTAATTAATTAATTAATTAATTAATTAAT 7942
Qy 610 caacagattctttcaattagagagtattcaaaacacactacaactactcgttgcacatt 669
Db 7943 AATAATATATTCACATTAATTTAATAGTTTATATATAATTAATTAATTTCTTTATTA 8002
Qy 670 ctatcactgatataaaagtgaaataaaatttaaaacacttagtttaataagaagaatt 729
Db 8003 TAATTAATTAGTACACATTTATTAATAATAATAATAATAATAATAATAATAATAATA 8062
Qy 730 ttattaaagcttgatcaataaatttaattctgataataaaataactatttaaacattaca 789
Db 8063 TATTATATATATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 8120
Qy 790 atgcttaataagattattatttatttatttatttatttatttatttatttatttatttatt 849
Db 8121 ATATTAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8180
Qy 850 ataaaaataacttactcttattttttatcactgataattatttaataataataact 909
Db 8181 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 8240
Qy 910 cccaaataaactattgcaaggttatgggaagatgaatttttactgttttttcatag 969

```







```
repeat_unit      650. .1022
                  /note="repeat I-A"
                  /rpt_type=tandem
repeat_unit      1023. .1360
                  /note="repeat I-B1"
                  /rpt_type=tandem
repeat_unit      1361. .1705
                  /note="repeat I-C/A"
                  /rpt_type=tandem
repeat_unit      1706. .2043
                  /note="repeat I-B2"
                  /rpt_type=tandem
repeat_unit      2044. .2388
                  /note="repeat I-C"
                  /rpt_type=tandem
misc_feature      2491. .2511
                  /note="deoxythymidylate stretch"
repeat_unit      2512. .2648
                  /partial
repeat_unit      2649. .3112
                  /note="repeat II-A"
                  /rpt_type=tandem
repeat_unit      3113. .3576
                  /note="repeat II-B1"
                  /rpt_type=tandem
repeat_unit      3577. .4040
                  /note="repeat II-B2"
                  /rpt_type=tandem
repeat_unit      4041. .4504
                  /note="repeat II-C"
                  /rpt_type=tandem
misc_feature      complement(4565..4585)
                  /note="deoxythymidylate stretch"
BASE COUNT      2271 a 131 c 74 g 2125 t
ORIGIN
```

```
Query Match      2.3%; Score 123.6; DB 33; Length 4601;
Best Local Similarity 43.6%; Pred. No. 3.3e+06;
Matches 1547; Conservative 0; Mismatches 1914; Indels 85; Gaps 19;

QY 494 accatcacctcttcatagtatcttctaagtaacaagaataaacttcataccgtaca 553
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3751 ATTAATTATTAATTAGTATAGTCTTTTAAATAAATAAATAATTTTAAATAAT 3692

QY 554 ttataactgattaaataaataaactattaatttgagcaaaaataattctcattcaac 613
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3691 TTTTAAATAAATGAAATAAATAAATAAATAATTTCAATTATAAATAATTTATTAA 3632

QY 614 agattctttcattagagagattccaaacacactcaactactgcttgcaactttctat 673
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3631 ATTTTGTGTTATTTTAAATAAACAATGATTTTATTATATAAATAATTTTATAAATA 3572

QY 674 cactgatatataaagtgaaataatttaaaactttagttttaatagaagaattttat 733
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3571 ATACATTTAGAATAATTTTAAATAAATTTATTTAAATTTTAAATAATTTATTTCT 3512

QY 734 taaagctttgaaatcaatttaactgataataaataactattaacaacttaacaatgc 793
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3511 ATATATATATATATATATAAATAATTCATTAATATAAATAATTTATAAATAATAA 3452

QY 794 ttaattaaggtattattattacctaatttcataacccttttataacaatttcataata 853
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3451 TTAATTAATTCATTTATTTATTTATAAATAAATAAATAAATAAATAAATAATGAGAATA 3392

QY 854 aagactctactcttatttttctactgataatttaataaataactataaactccca 913
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3391 TAAATTTTATAAATATATCTACATTTTAAATTTTAAATTTTAAATTTTAAATTTATTA 3332

QY 914 aataaactattgcaaggttatggtaattgaaatttttactgttttttctcattagttt 973
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3331 GATATATAATAATATATAATTTTATATATATAAATAATCTATTAATTTTATAATTAG 3272
```

```
QY 974 cttaacaatagccaatcatcgcttttctcctttaacattaaagttacacatgaaaaataga 1033
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3271 TATATAGTTTTTTTTTAAAAAAAATAATTTATTTTTTAAAAAATTTTAAAAAATTTGA 3212

QY 1034 taatggaatgga--agttacgtgattcccaaatcatcgcgcaccagcagtcacatgcacatg 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3211 AAAATAAATAAATAATATATTTTCATTATATAAATAATTTATTTATTAATAAATTTTGTGTTATTT 3152

QY 1092 gtattatacaaaagtcggtaggaactgatgccagtaggatactctggttagcacattttt 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3151 TTAAAAAACATGATTTTATATATAAATAATTTTATATAAATAATTAACATTTTAAAGAAAT 3092

QY 1152 ttggaacacttaagtgttagtggaacagaaaaatttccctaactcatcgcacacacttagt 1211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3091 TTTAAAAAATTTATATAAATAATTTTAAATAATTTTAAATTTTCTATATATATATATATAT 3032

QY 1212 aatataggcggaatttcaatgcaagcacacatctcaatttcttactatactactcagaatta 1271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3031 TATATAA-----ATATTCAATAATATATAAATAATTTATAAATAATATAAATAATTAAT 2982

QY 1272 ataccaaaacaataatttctcttgcgaatggatatatgaatcacagacagaatgcagaatttt 1331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2981 AAATTATTTATTTATTTATATAAATAATTTAAATAAATAAATAATTAAGAAATATAAAT 2922

QY 1332 aagggttaccgcacaagcattataagaagaaacaaaggttagtctttagaagaagaaaaatg 1391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2921 TTTTAAATAATTAATCTACATTTTAAATTTTAAATTTTAAATTTTATTAATTTTAGATATA 2862

QY 1392 agagttgaaagccaagcaaaaaacatactactagaagaagaatggaaaaatgcattttttac 1451
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2861 TAATAATATATTAATAATTTATATATA-TATAAATATCTATTAAATTTATAATTTAGTATAT 2803

QY 1452 aatggatatggcagaccagtagtgatgggaacaatgaattagcaactacaacaagaa 1511
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2802 AGTTTTTTTTTAAAAAAAATAATTTTAAAAATAATTTTAAAAATAATTTTAAAAATAAT 2743

QY 1512 gttgctgaagcctttcataagctacatttagtccctaataatgctctatttaattgttaact 1571
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2742 AATAAATTTATTTTCATTATTAATAAATTTATTTTAAATAATTTTGTGTTATTTTAA 2683

QY 1572 ggagatgcagatcccaagaagtaatcacactgcgaacaacatactatgggaaaaatacaca 1631
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2682 AAACATGATTTTATTTATATAAATAATTTTATAAAAAATAATACATTTAAGAAATTTTAA 2623

QY 1632 tctaaataaagaacaccttcaagtcagtttaggttagaaccaccgcacataaacaataatg 1691
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2622 AAAATTTATTTAAATTTATTTAAATAATTTAAATTT-----TTCATATATATA 2575

QY 1692 actttaacattaaaagacagcttcagtagaaatcccagaaactgtttttaatgtatcaata 1751
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2574 TATATATTAATAATTTCAATATATATAAATTTATAAATTTATATAATTAATTAATTAAT 2515

QY 1752 ccaatgtgtattaccataaaaaactacatacttaactgatgttagcagaaaaactcoggt 1811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2514 ATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2455

QY 1812 agtggtaaatccagctgcttttaactgatgtttgtaattacaactccaatagttacatcg 1871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2454 AAATTTTATTTAAATTTAGTTTATTA-AAATTTATTTATTTAAATTTAAATTTAAATG 2396

QY 1872 ataaaaacagattataacttaactgacagcgataattacacctttccattgtagctata 1931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2395 AATAGATTTTATTTAATTTATTAATTTATTAATTTATTTATTTAATAAATTTAAATTTAT 2336

QY 1932 cctaaaaacgggactctctacagagagctgtagaacaagaataatcotaatgataaataat 1991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2335 TTTCAATTTAATATATATATATATATATAAATAAATAAATAAATAAATAAATAAATAAAT 2276

QY 1992 tatttgagaataatgggaatttcagcagaaata---tttgagaagtgcaagataaagtaaaa 2048
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2275 TTTTATTAAAAATAATTTATTTATAAATAATAGTTTATTAAAGTATAAATTTAATTAATTAAT 2216
```

QY 2049 gacatttaactatgcatttgacaggactaaacttccatactatattttat-----gg 2101  
 Db 2215 TTTTAAAAAATAATTTTAAAGTTTAAATATATAATAATTTTAAATAGGGG 2156  
 QY 2102 catgatcctaactaggagtcacgactcagaaatcagtaatttaccagatcaccatag- 2160  
 Db 2155 GAATAAATTTATTTCAATTTTACATATATATATATATATATATATATATATATAT 2096  
 QY 2161 aceaagtagtattccaaagtgttaactccgctatggaaaatctcttccaaacaataaa 2220  
 Db 2095 TCAGATTTAGTCATTAATAATAAATTTATTTATTAATTTATTAATTTAAATGAAATTA 2036  
 QY 2221 gattaacggcgcatattaccctaattgaggaatagttatgagaaacatatgttggtacac 2280  
 Db 2035 AATTATATGATATATATAATAATATGAATTTTAAATTTTAAATTTTAAATTTTAAAT 1976  
 QY 2281 attaatattgattcttcttccatcaatcatatgcaatgactcacaatttaacataaa 2340  
 Db 1975 TTTTATATATTAATAATAATTTTATTAATTTTAAATAATTTTATTTTAAATAAT 1916  
 QY 2341 agaagctcaacta-----aaaataaaatcacactatctatgttggaacatcata 2390  
 Db 1915 TTATTTAAAAATAGTTTATTAAGTATATTTAATAATTTTAAATTTTAAATTTTAAAT 1856  
 QY 2391 acctaccacaatttcccttaaaatttgcatccaagaagcagatagcgttatgatgct 2450  
 Db 1855 TATTTTAAAGTTTAAATTTATACAAATAATTTATGAATAGGGGAATAAATTTATTTTCA 1796  
 QY 2451 ttgataagcaagacttgcatcttaccatcaaaatatttaaacgaagatcaaaaaaca 2510  
 Db 1795 TTTTATATATATATATATATATATATATATATATTTTCAATTTTATTAATAATAA 1736  
 QY 2511 actatgctctcagtttgacaaa--caattagaaggcaaggtatagacttaaaatttga 2568  
 Db 1735 ATATTTTATATATATATATATATTTAATTTAATTTAATAATTTATTAATAATAATA 1676  
 QY 2569 tatagacctt-----gacatttttatatcatcatttcaaatcattatctgcctgaattca 2624  
 Db 1675 TATAAATAAATTTGAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1616  
 QY 2625 aagccctagtttactcagtttgcatattcaacacgcgcagatcaagaataattca 2684  
 Db 1615 TCTTATTAATTTTAAATAATTTTATTAATAATAATTTTAAATAATAATTTATTAATAATA 1556  
 QY 2685 atagataatagcagaacagatgcatcatttcaatcattatcttgcctgaattca 2744  
 Db 1555 TAAGTATATTAATAATAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTA 1496  
 QY 2745 t-----agctacacagaaatgaatcagctatattcaaggagcccatattctaac 2797  
 Db 1495 TATAATAAATTTATGAATAGGGGAATAAATTTATTTTCAATTTTACATATATATATATAT 1436  
 QY 2798 aaagtctcaggagacatttaatacaatcaataatacaacacgcgcagagagcgttcattatat 2857  
 Db 1435 ATATATATATACAAATTAATTTTCAATTTTATTAATAATAATTTTATTAATAATAATA 1376  
 QY 2858 ataaaaaatagtttgacagaagcaaaatcggtatcagcgacgcagagagatgtagatcca 2917  
 Db 1375 TTATATATTTTAAATTTGAAATTTAATTTATATATATATATATATATATATATATATAT 1316  
 QY 2918 acacagctatcaaaatttactagataaaatattcttccaaattgccaatctggtgaataac 2977  
 Db 1315 TTTTATAA--AAATCAATTTTAAATTTTATATATATTAATAATAATTTTATTAATAATTT 1259  
 QY 2978 aaaaataccataccagatcagactgtaataagagaagacacacattattatatgacagaga 3037  
 Db 1258 TAAAAATAATTTTATTAATAATAATTTTATTAATAATAATTTTATTAATAATAATAATA 1199  
 QY 3038 gatgc-----accacaagtgctcaatgttctgctacagacacagctaccata 3084  
 Db 1198 TAAATCAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1139  
 QY 3085 tcacagcaagactatcatgcatcaactgttctcaactatgctaggcgagattagctct 3144

Db 1138 GAATAGGGGAATAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1079  
 QY 3145 caattcaatatttaattagatagaatttaagacacaggttaggattaaacattacacacagtagcg 3204  
 Db 1078 TATTTTCAGATTTAGTGATTAATAATAATTTTATTTTATTTTATTTTATTTTATTTTAA 1019  
 QY 3205 ttcatatctaacaatgaatcatagtaagtgtctatttgggacaaatattcactgataaatc 3264  
 Db 1018 ATTAATAATTTAT 959  
 QY 3265 caca---gtaacaaaatgtatatccgctcttcaacagattatttagagcacattaaaaagta 3321  
 Db 958 TTAATTTTATTTATTAATAATAATAATTTTCTTATTAATTTTAAATAATTTTATTTTATA 899  
 QY 3322 tggagtgtgatgaagcacacttttgcatt-tgcaaaatcttagttaccactctttttttt 3380  
 Db 898 AATAATTTTATTAATAATAATAGTTTATTAAGTATATTTAATAATAATTTTATTTTAA 839  
 QY 3381 tatctatgttaaatcaacaataatgttttagtgagatattgttaagcttacaattcacogate 3440  
 Db 838 AAAAAATACTTTTAAAGTTTAAATTTATATATATATATATATATATATATATATATAT 779  
 QY 3441 tagatccgagtttatattaataacaattcttactacaagaacatacaatacaatagagaag 3500  
 Db 778 TTTTATTTATTTATGATA 720  
 QY 3501 taataataaattgccagaataattttctcaatgaatctagtaaatctgaagtaggaaaaa 3560  
 Db 719 TATATATAATTTTTCAGATTTAGTGATTAATAATAATTTTATTTTATTTATTTATTCG 660  
 QY 3561 acaataacaataatggcaacaataatagatgctcaaaaaacacacacactctggttaagtatac 3620  
 Db 659 TAATTTAATTAATAATCAAAATTTATTTTATATATATATATATATTTTAAATAAATTTAT 600  
 QY 3621 aggttattgtatttact-acaagttattctttaggttatttaagtaagataaagttagct 3679  
 Db 599 TACTTTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTCA 540  
 QY 3680 tcaatcaataaaaaaacatttaaccaaggtgttagctctcacggagagctctattataag 3739  
 Db 539 TTTTGTGTAATAATAATAATTTTATTAATAATAATTTTATTTTATTTTAAATAAATTTT 480  
 QY 3740 ctttaacctgggataataatgaagttttgtcattgttaagcaaaaaatttagtaatac-- 3797  
 Db 479 TTTAAATTTTATATTTATATATATTTTAAATAAATTTGTAATAAATAAAGTTTAAATAACT 420  
 QY 3798 -aatatcaaaatttcttctacaggattatttgcacctaccatacaacatttatattaga 3856  
 Db 419 AAAAAATAAATTTTATTTAGTCATTTAAAGAAATTAATCTATAATTTTAAATAATTTT 360  
 QY 3857 aatgacacagatcacacacatacaataaatattcactacaattcaatttaataaacaatg 3916  
 Db 359 TTAATAATAATAAATTTTAAATAATAAATTTTAAATAATAAATTTTAAATAATAATTTT 300  
 QY 3917 agtatttttacttaattatttatttttttttttttttttttttttttttttttttttttt 3976  
 Db 299 AAGTGTAAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTATTTTATTTATTTGAT 240  
 QY 3977 caataa 3982  
 Db 239 AAATTA 234

RESULT 15  
 DMU37541/c  
 LOCUS DMU37541 19517 bp DNA circular INV 04-APR-2000  
 DEFINITION Drosophila melanogaster complete mitochondrial genome.  
 ACCESSION U37541  
 VERSION U37541.1 GI:1166529  
 KEYWORDS  
 SOURCE Drosophila melanogaster.  
 ORGANISM Mitochondrion Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 12511 to 12682)  
Clary,D.O., Goddard,J.M., Martin,S.C., Fauron,C.M. and Wolstenholme,D.R.  
Drosophila mitochondrial DNA: a novel gene order  
Nucleic Acids Res. 10 (21), 6619-6637 (1982)

2 (bases 5269 to 5695)  
Clary,D.O., Wahleithner,J.A. and Wolstenholme,D.R.  
Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequences and comparisons to mammalian mitochondrial RNA genes  
Nucleic Acids Res. 11 (8), 2411-2425 (1983)

83220794  
Drosophila melanogaster mitochondrial DNA, a novel organization and genetic code  
Nature 304 (5923), 234-241 (1983)

83245048  
Satta,Y., Ishiwa,H. and Chigusa,S.I.  
Analysis of nucleotide substitutions of mitochondrial DNAs in Drosophila melanogaster and its sibling species  
Mol. Biol. Evol. 4 (6), 638-650 (1987)

88174373  
Garesse,R.  
Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary considerations  
Genetics 118 (4), 649-663 (1988)

88212147  
6 (bases 441 to 2967)  
Satta,Y. and Takanata,N.  
Evolution of Drosophila mitochondrial DNA and the history of the melanogaster subgroup  
Proc. Natl. Acad. Sci. U.S.A. 87 (24), 9558-9562 (1990)

91088557  
7 (bases 14215 to 14512)  
Ballard,J.W., Olsen,G.J., Faith,D.P., Odgers,W.A., Rowell,D.M. and Atkinson,P.W.  
Evidence from 12S ribosomal RNA sequences that onychophorans are modified arthropods  
Science 258 (5086), 1345-1348 (1992)

93088057  
8 (bases 14917 to 19517)  
Lewis,D.L., Farr,C.L., Farquhar,A.L. and Kaguni,L.S.  
Sequence, organization, and evolution of the A+T region of Drosophila melanogaster mitochondrial DNA  
Mol. Biol. Evol. 11 (3), 523-538 (1994)

94285822  
9 (bases 1 to 408; 13319 to 19517)  
Lewis,D.L., Farr,C.L. and Kaguni,L.S.  
Drosophila melanogaster mitochondrial DNA: completion of the nucleotide sequence and evolutionary comparisons  
Insect Mol. Biol. 4 (4), 263-278 (1995)

96423163  
10 (bases 1 to 19517)  
Lewis,D.L., Farr,C.L. and Kaguni,L.S.  
Direct Submission  
Submitted (03-OCT-1995) Laurie S. Kaguni, Biochemistry Department, Michigan State University, East Lansing, MI 48824-1319, USA  
Location/Qualifiers  
1. .19517  
/organism="Drosophila melanogaster"  
/organelle="mitochondrion"  
/db\_xref="taxon:7227"  
/note="derived from new and previously submitted sequences; sequence is a composite containing sequences obtained from different Drosophila melanogaster strains"

1. .65

trna

/product="trna-ile"  
complement(97. .165)  
/product="trna-gln"  
171. .239  
/product="trna-phe"  
240. .1265  
/codon\_start=1  
/transl\_table=5  
/product="NADH dehydrogenase subunit 2"  
/protein\_id="AAC47811.1"  
/db\_xref="GI:1166530"  
/translation="MFNNSKILFITIMIGIILITVTSNWLGAWGLINLLSFPL  
LSDNNLMTESLKYFLQVLASTVLFSSILLMLKNNMNEINFSMIIMSALL  
LKSQAAPFFHFWFNMMEGLTWMALMTWOKIAPMLISYLNKYLISVLSVII  
GAIGLNQTSRLKLMFASINHLGWLSSLMISEIWLILFFYSFLSVLTFMFNIF  
KLFLNQLFSWFVNSKILKFTLFMNFSLGGLPPLFGPLKVLVIOQLCLCNOYFULT  
IMMSTLITLFFYLIRICYSAFMNYFENNWMKMNNSINYNMYMTITFSIFGLFLI  
SLYFMP"  
1264. .1329  
/product="trna-trp"  
complement(1322. .1383)  
/product="trna-cys"  
complement(1403. .1468)  
/product="trna-tyr"  
join(1470. .1472,1474. .3009)  
/codon\_start=1  
/exception="mechanism underlying reading frame shift after first codon uncertain"  
/transl\_table=5  
/product="cytochrome c oxidase subunit I"  
/protein\_id="AAC47812.2"  
/db\_xref="GI:7412849"  
/translation="MSROWLSTNKHDKIGTYLFGAWAGMVGTSILIRAEGLHPG  
ALIGDQIYNVIVTAFAFIEMFMVPMIGGFGNWLVLMLGAPDMPRMNMSEFW  
LPPALSLLVSMWENGAGTGWYPPLSAGIAHGASVDLAFSLHLAGISSILGA  
LVFTIVINRSTGISLDRMPLFVSWVITALLLSLPLAGAITMLTDRNLNFSF  
FPPAGGDPLYOHLFFHFGPEVYILIPFGMISHIISQESKKETFGSLGMIYAF  
LAIGLGFVWAHMFVTGMDVDTRAYFTSATMIAVPTGKIFSLATLHGTQLYS  
PAILWALGFVLTGGLGVLANSSVDIILHDYTVVVAHFVYLSMGAVPDAITWYN  
IHWPPLTGLTLNNKWLKSHFIMFVGNLTFFPQHFGLAGMPRYSYDPDAYTTWN  
IVSTIGSTISLLGLFFFIWESLSVSRQVYVPIQLNSSIEWYQNTPTAEHSYSELP  
LUTN"  
3012. .3077  
/product="trna-leu"  
3083. .3767  
/note="TAA stop codon is completed by the addition of 3' A residues to the mRNA"  
/codon\_start=1  
/transl\_except=(pos:3767,aa:TERM)  
/transl\_table=5  
/product="cytochrome c oxidase subunit II"  
/protein\_id="AAC47813.1"  
/db\_xref="GI:1166532"  
/translation="MSTWANLQDSASPLMEQLIFFHDBALLILYMTVLYGLMFM  
LFNNYVNRFLHGOIEMITLPAILLFALPSLRVLLDEINEPSVTSKLSIGH  
QWYSEYSDFNNEIDSYMTNMLDTGFLDDVNRVLPVPMNSOIRILVTAADV  
HSWTVPALGVKVDGTPGRNLQTNFFINRPGLYGQCSEICGANHSFMPVIESVPVNY  
FIKWISSNNS"  
3768. .3838  
/product="trna-lys"  
3840. .3906  
/product="trna-asn"  
3907. .4068  
/codon\_start=1  
/transl\_table=5  
/product="Atpase 8"  
/protein\_id="AAC47814.1"  
/db\_xref="GI:1166533"  
/translation="MPQMAPISWLLLIIFSIITFILFCSINYSYMSPKSNELKNI  
NLNSMNWKW"  
4062. .4736  
/codon\_start=1  
/transl\_table=5





Search completed: December 3, 2000, 00:11:50  
Job time: 33425 sec

---

THIS PAGE BLANK (uspto)